

# High-order boundary integral equation solution of high frequency wave scattering from obstacles in an unbounded linearly stratified medium

Alex H. Barnett<sup>a</sup>, Bradley J. Nelson<sup>b</sup>, J. Matthew Mahoney<sup>c</sup>

<sup>a</sup>Department of Mathematics, Dartmouth College, Hanover, NH, 03755

<sup>b</sup>Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, 94305

<sup>c</sup>Department of Neurological Sciences, University of Vermont, Burlington, VT, 05405

---

## Abstract

We apply boundary integral equations for the first time to the two-dimensional scattering of time-harmonic waves from a smooth obstacle embedded in a continuously-graded unbounded medium. In the case we solve the square of the wavenumber (refractive index) varies linearly in one coordinate, i.e.  $(\Delta + E + x_2)u(x_1, x_2) = 0$  where  $E$  is a constant; this models quantum particles of fixed energy in a uniform gravitational field, and has broader applications to stratified media in acoustics, optics and seismology. We evaluate the fundamental solution efficiently with exponential accuracy via numerical saddle-point integration, using the truncated trapezoid rule with typically  $10^2$  nodes, with an effort that is *independent* of the frequency parameter  $E$ . By combining with high-order Nyström quadrature, we are able to solve the scattering from obstacles 50 wavelengths across to 11 digits of accuracy in under a minute on a desktop or laptop.

**Keywords:** scattering, acoustic, Helmholtz, graded-index, refraction, gravity, quantum, integral equation

**2010 MSC:** 65N38, 65N80, 34M60, 65D20

---

## 1. Introduction

Problems involving time-harmonic waves in media whose wave speed or refractive index varies continuously in a layered fashion are common in both the natural and engineered worlds. In acoustics, underwater sound propagation [1, 2], and environmental noise modeling in the presence of a thermal gradient [3] both involve continuously stratified wave speeds. In electromagnetics, continuously stratified media occur in ionospheric propagation [4] and nano-scale optical devices (see [5] and references within). In elastodynamics, similar models play important roles in seismology since wave speed grows in a piecewise continuous fashion with depth into the earth [6, Sec. 2.5.3], and in designing functionally graded materials [7]. In quantum physics the same equations as in acoustics arise when gravitational or electric fields influence the motion of fixed energy particles [8]. In each case, when the varying medium is acoustically large (many wavelengths across), or unbounded, accurate numerical solution of wave propagation and scattering remains challenging.

We will solve the following scalar-wave exterior boundary value problem (BVP), where  $\Omega \subset \mathbb{R}^2$  is a given bounded obstacle with smooth boundary  $\partial\Omega$ , and  $f$  is smooth Dirichlet data on  $\partial\Omega$ ,

$$(\Delta + k(x_2)^2)u(x_1, x_2) = 0 \quad \mathbf{x} := (x_1, x_2) \in \mathbb{R}^2 \setminus \overline{\Omega}, \quad (1)$$

$$u = f \quad \text{on } \partial\Omega, \quad (2)$$

where  $\Delta := \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$  is the Laplace operator, with the specific vertical wavenumber variation  $k(x_2)$  given by

$$k(x_2)^2 = E + x_2, \quad (3)$$

and outgoing radiation conditions for  $u$ . The latter, given in Definition 1, are required for uniqueness of the solution. In applications the potential  $u$  represents pressure, wavefunction, or a component of electric or magnetic field.

The general relationship  $k = \omega/c$ , where  $\omega$  is frequency and  $c$  wave speed, means that in the frequency-domain (fixed  $\omega$ ) case,  $k$  is proportional to the *refractive index* and inversely proportional to the wave speed. In (3) the inverse square of wave speed (sometimes called *sloth*) is linear in the vertical ( $x_2$ ) coordinate, a model found in seismology [6,

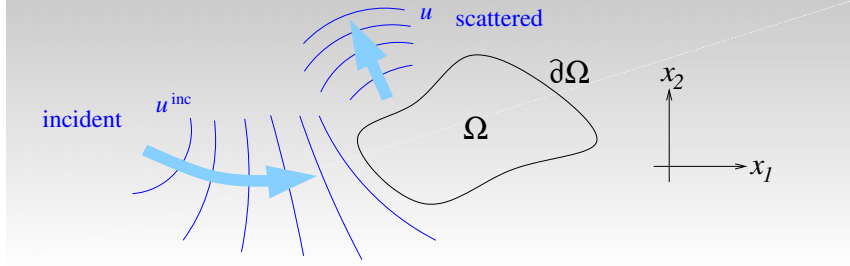


Figure 1: Geometry for the scattering problem embedded in a stratified medium. Wave speed decreases (refractive index increases) in the vertical  $x_2$  direction.

Sec. 2.5.2.2]; in the electromagnetic case (3) corresponds to linear variation in permittivity [9, Sec. 2.5.1]. Recalling that the Helmholtz equation  $(\Delta + E)u = 0$  models free-space quantum particles at energy  $E$ , we call (1) with (3) the “gravity Helmholtz equation” because it is a non-dimensionalized<sup>1</sup> model for quantum particles at energy  $E$  in a uniform gravitational [8] or electric [10] field, i.e. a linear potential. Its one-dimensional (1D) solution is the Airy function, and its application goes back at least to Hartree’s 1931 work on the ionosphere [4, Sec. 6]. The constant  $E$  sets the square of the wavenumber at the height  $x_2 = 0$ ; the waves have evanescent (modified Helmholtz) character for  $x_2 < -E$ , changing to oscillatory (Helmholtz) character for  $x_2 > -E$ . The asymptotic behavior of solutions to (1) is radically different in the horizontal and vertical directions, with waves eventually “dragged” into a narrow upwards-propagating beam; see Fig. 2(b). In the optical and acoustic setting, the imaginary refractive index for  $x_2 < -E$  could be relevant for graded metamaterials, although a more common application of the efficient PDE solver we present might be to acoustic or electromagnetic propagation in subregions of the plane (a half-space, etc).

In the usual setting of scattering theory (see Fig. 1) an incident wave  $u^{\text{inc}}$  satisfying (1) in the entire plane impinges on the obstacle; the scattered wave is then  $u$ , the solution to the above exterior BVP with boundary data  $f = -u^{\text{inc}}$  on  $\partial\Omega$ . The physical potential is then  $u^{\text{inc}} + u$ . The Dirichlet case we study corresponds to sound-soft acoustics, or  $z$ -invariant Maxwell’s equations with a perfect electric conductor in transverse-magnetic polarization. The Neumann (sound-hard) case can be solved with similar tools [11]. We will also solve the interior Dirichlet BVP, with applications to graded-index optics, and to transverse acoustic or optical modes in a bending waveguide in 3D approximated by an “equivalent profile” in which the square of refractive index varies linearly [12].

We propose boundary integral equations (BIE) as an efficient and accurate numerical method to solve (1)–(2). This demands being able to compute values and first derivatives of  $\Phi(\mathbf{x}, \mathbf{y})$ , the *fundamental solution* to (1), where  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$  are target and source points respectively. Recall the definition that, for a source point  $\mathbf{y} \in \mathbb{R}^2$ ,  $\Phi(\cdot, \mathbf{y})$  is the radiative solution to the PDE

$$-(\Delta_{\mathbf{x}} + k(x_2)^2)\Phi(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), \quad (4)$$

where  $\delta$  is the Dirac delta distribution in  $\mathbb{R}^2$ . In contrast to the common situation,  $\Phi$  is no longer an elementary or special function of distance  $|\mathbf{x} - \mathbf{y}|$ ; this is clear in Fig. 2. A large part of our contribution is an efficient numerical method for evaluation of  $\Phi$ , by applying quadrature to the Fourier transform of an analytical solution to the time-dependent Schrödinger equation in a linear potential [13]. Unfortunately the integral is highly oscillatory, especially as  $E$  grows, thus we use deformation of the contour into the complex plane, passing through the saddle (stationary phase) points and using the trapezoid rule [15] to achieve exponential accuracy with effort independent of  $E$ . The saddle points will have an elegant interpretation as the classical ray travel times. The cost of each evaluation of  $\Phi$  is only a few hundred complex exponential evaluations, hence we achieve typically  $10^5$  evaluations per second.

### 1.1. Relation to previous work on frequency domain wave propagation in layered media

Accurate numerical propagation of high frequency waves in a variable medium is numerically challenging: conventional “volume” discretization methods such as finite differencing (FD) [16] and finite elements (FEM) require several degrees of freedom per wavelength to achieve reasonable accuracy; moreover, in order to avoid “pollution

<sup>1</sup>We chose a unity constant in front of  $x_2$  without loss of generality since adjusting this constant is equivalent to rescaling the domain  $\Omega$ .

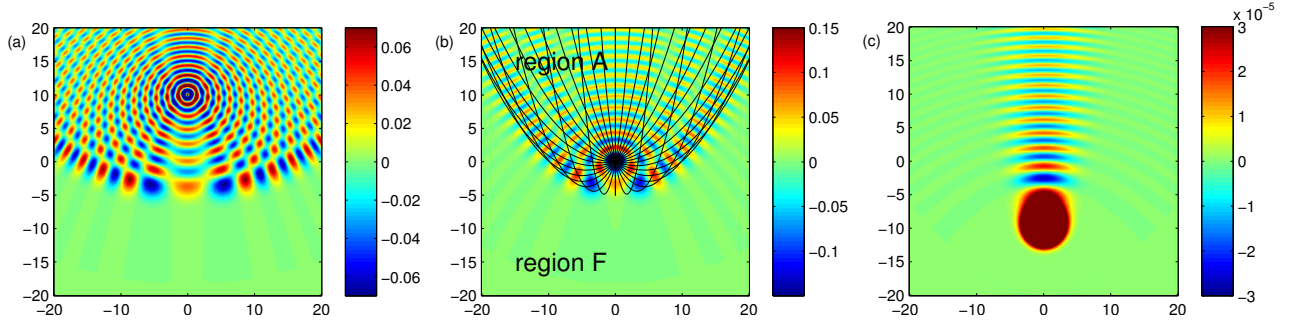


Figure 2: Real part of fundamental solution  $\Phi(\cdot, \mathbf{y})$  plotted in  $\mathbb{R}^2$  for the case  $E = 5$  and three choices of source location  $\mathbf{y}$ : (a)  $\mathbf{y} = (0, 10)$ ; (b)  $\mathbf{y} = (0, 0)$ ; (c)  $\mathbf{y} = (0, -10)$ . In (b) we also show parabolic classical ray trajectories emanating from the source  $\mathbf{y}$ , which themselves all lie within region A (a parabola with focus  $\mathbf{y}$ ), discussed in Sec. 2.1. Notice the color scale in (c) indicating the very small amplitude of the propagating beam.

errors” the degrees of freedom per wavelength must grow with frequency [17]. The resulting linear systems are so huge that iterative solvers are almost always used, and yet preconditioning has mostly been unsuccessful for the high-frequency Helmholtz equation, especially for high-order discretizations, and is a topic of current research [18]. The radiation condition must still be approximated via artificial absorbing boundary conditions (e.g. perfectly matched layers) [19] [9, Sec. 4.7].

At high frequencies, ray approximation is useful [6] and geometric diffraction theory can approximate the interaction with simple obstacles. However, such approximations break down at turning points (such as at  $x_2 = -E$ ) and for geometric details on the wavelength scale. Parabolic approximation (i.e. one-way wave equation) methods [20] handle only a limited range of propagation directions, and cannot account for back reflections. Several of these methods are reviewed in the underwater acoustic and elastic contexts in [21].

When the medium (PDE coefficient) is piecewise constant, reformulation as a boundary integral equation (BIE) [22, Ch. 3] [23] [9, Ch. 8] is popular due to several advantages:

- the unknowns live on the boundary (or material interfaces) rather than the volume; this reduction in dimension by one greatly reduces the number of unknowns  $N$ , especially at high frequencies, and simplifies the geometric issues (meshing, etc);
- when a second-kind formulation is used, it remains *well-conditioned* (and hence iterative methods rapidly convergent) independent of the number of discretization nodes used;
- radiation conditions are already built into the representation and need not be enforced, unlike in FD or FEM;
- fast algorithms such as the fast multipole method [24] or fast direct solvers [25, 26] can reduce the solution time to  $O(N)$  for low frequencies;
- in the two-dimensional case, high-order quadratures on boundary curves are easy to implement [27, 28].

This has enabled the scattering from objects (in a uniform medium) thousands of wavelengths across to be solved efficiently to many digits of accuracy (e.g. see [29]).

In contrast, we care about scattering in a continuously-varying medium. If this medium were constant outside a bounded region, a Lippmann–Schwinger (volume integral) equation [22, Ch. 8] could be used, or coupling of direct discretization methods to BIE [30, 31]. Tools also exist for BIE within media with a finite number of constant layers [32]. The method of the present paper extends the above advantages of BIEs to a particular problem where the stratified medium variation—and the resulting wave propagation—is smooth and *unbounded* in all directions. We are not aware of previous applications of BIE to such a case. The only similar work we know of is that of Premat–Gabillet in their environmental acoustics code Meteo-BEM [3], who use BIEs with the Green’s function for a linear wave speed profile. However, they approximate the Green’s function using a discrete sum over 1D eigenfunctions, an approach that works only when waves are trapped by a ground plane; this would fail in the case of unbounded propagation. Also, since their BIE is of Fredholm first kind, the convergence rate of an iterative solver would be poor.

**Remark 1.** Our approach to evaluate the Green's function is reminiscent of the Sommerfeld integral (spectral representation) commonly used for layered media [9, Ch. 2], [32]. Yet, although both methods exploit numerical quadrature of a contour integral, they are distinct, with crucial differences. In the Sommerfeld approach the integration variable is a transverse wavenumber, and a vertical ODE has to be solved for each contour quadrature node; for the profile (3) this would demand Airy functions. The number of quadrature nodes needed grows linearly with wavenumber, for fixed source-target separation. In addition, the decay of the Sommerfeld integrand is known to be very slow when the vertical separation is small, demanding various windowing approximations [32]. In contrast, in our proposed scheme the integration variable represents time, the integrand involves only exponentials, and by choosing appropriate complex contours the number of nodes is independent of wavenumber. Of course, the Sommerfeld approach has the advantage over our scheme that, assuming the ODEs can be solved fast enough, arbitrary profiles  $k(x_2)$  could be handled.

### 1.2. Outline of the paper

We use the remainder of the introduction to state a radiation condition that allows a unique solution to our BVP (this is proved in Appendix A). In Sec. 2 we present an integral formula for the fundamental solution (4) for the PDE (1); here the radiation condition derives from causality in the time domain. We then use potential theory to reformulate the BVP as an integral equation on  $\partial\Omega$  in Sec. 3, and present its high-order numerical solution, which demands many evaluations of the fundamental solution. Sec. 4 is the key part of the paper in which we present efficient new contour quadrature algorithms for this task. In Sec. 5 we present numerical tests of convergence and speed for both the interior and exterior BVPs. We draw some conclusions and discuss future work in Sec. 6.

### 1.3. The radiation condition for the BVP

Recall that for the constant- $k$  Helmholtz equation  $(\Delta + k^2)u = 0$  in  $\mathbb{R}^2$ , the Sommerfeld radiation condition [22, (3.62)] is  $\partial u / \partial r - iku = o(r^{-1/2})$ , holding uniformly in angle, where  $r := |\mathbf{x}|$ . This corresponds to outgoing waves at infinity. It guarantees a unique solution to exterior BVPs [22, Sec. 3.2], for instance the case of Dirichlet data (2). Radiation conditions are also known for stratified media that are eventually constant or tend to a constant in upper and lower half-planes [33, 34], for variable media that tend towards a constant at large distances [35], and for scattering from unbounded rough surfaces in a uniform medium [36]. For our exterior gravity Helmholtz equation there are no trapped waveguide modes because the refractive index is monotonic in  $x_2$ , simplifying the situation from that of [33, 34]. And yet, we have not been able to find a radiation condition in the literature that applies in our case where the wavenumber is unbounded in one direction.

Hence we propose the following new radiation condition, recalling the notation  $\mathbf{x} = (x_1, x_2)$ .

**Definition 1** (Radiation condition). *A solution  $u$  to (1) in the exterior of a bounded domain  $\Omega \subset \mathbb{R}^2$ , with medium defined by (3), is called radiative if*

$$\lim_{x_2 \rightarrow +\infty} \frac{1}{k(x_2)} \int_{-\infty}^{\infty} \left| \frac{\partial u}{\partial x_2} - ik(x_2)u \right|^2 dx_1 = 0 \quad (5)$$

$$\lim_{x_2 \rightarrow -\infty} \int_{-\infty}^{\infty} |u|^2 + \left| \frac{\partial u}{\partial x_2} \right|^2 dx_1 = 0 \quad (6)$$

$$\lim_{L \rightarrow \infty} \lim_{x_1 \rightarrow \pm\infty} \int_{-L}^L |u|^2 + \left| \frac{\partial u}{\partial x_1} \right|^2 dx_2 = 0 \quad (7)$$

The first condition states that the flux is eventually upwards-going on positive horizontal slices; the other two guarantee enough decay that the flux tends to zero on the sides and bottom of a large rectangular box. Note that these conditions could most likely be tightened; however, they are adequate for our purpose, namely to prove in Appendix A the following uniqueness result, analogous to [22, Thm. 3.7] for the Helmholtz equation. This places our BVP on a more rigorous footing.

**Theorem 1.** *There is either zero or one radiative exterior solution to (1)–(3).*

## 2. The fundamental solution and its ray interpretation

In this section we derive an integral formula for the fundamental solution for our PDE (1) in  $\mathbb{R}^2$ , and give some of its properties. In fact, since it requires no extra effort, we work in  $\mathbb{R}^n$  and then specialize to  $n = 2$ . Let  $\mathbf{x} = (\mathbf{x}', x_n)$  where  $\mathbf{x}' = (x_1, \dots, x_{n-1})$  is the transverse coordinate and  $x_n$  is the vertical one. The gravity Helmholtz equation in  $\mathbb{R}^n$  is  $(\Delta + E + x_n)u(\mathbf{x}) = 0$ . Recall that the fundamental solution is defined by (4). We will exploit causality in the time domain to obtain a solution with physically correct radiation conditions, so call this the “causal” fundamental solution (although see Remark 3).

**Lemma 1** (Bracher et al. [8]). *The causal fundamental solution to the gravity Helmholtz equation  $(\Delta + E + x_n)u(\mathbf{x}) = 0$  in  $\mathbb{R}^n$  with source point  $\mathbf{y} \in \mathbb{R}^n$  is given by*

$$\Phi(\mathbf{x}, \mathbf{y}) = \frac{i}{(4\pi i)^{n/2}} \int_0^\infty \frac{1}{t^{n/2}} \exp i \left[ \frac{|\mathbf{x} - \mathbf{y}|^2}{4t} + \left( \frac{x_n + y_n}{2} + E \right) t - \frac{1}{12} t^3 \right] dt. \quad (8)$$

Its proof exploits the fact that the time-dependent *Schrödinger equation* has an analytically known fundamental solution in a linear potential. We will show that the integral in (8) is in fact the Fourier transform from time  $t$  to energy  $E$ ; note that this is distinct from the more usual connection of frequency-domain fundamental solutions to the *wave equation*, for instance in the Cagniard–de Hoop method [9, Sec. 4.2].

For convenience we simplify and rephrase the derivation of Bracher et al. [8] in a more mathematical language, and in dimensionless units. Our definitions of the Fourier transform from time to energy will be, in terms of a general function  $f$ ,

$$\int_{-\infty}^\infty \tilde{f}(t) e^{iEt} dt = f(E), \quad \frac{1}{2\pi} \int_{-\infty}^\infty f(E) e^{-iEt} dE = \tilde{f}(t).$$

Similarly, our definition for spatial Fourier transforms is

$$\int_{\mathbb{R}^n} \hat{f}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k} = f(\mathbf{x}), \quad \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} = \hat{f}(\mathbf{k}).$$

We now prove the lemma.

*Proof.* We will isolate the last coordinate with the notation  $\mathbf{x} = (\mathbf{x}', x_n)$  and  $\mathbf{y} = (\mathbf{y}', y_n)$ . Suppressing for now the  $\mathbf{y}$  dependence, but making the dependence on  $E$  explicit, the fundamental solution obeys

$$(\Delta + E + x_n)\Phi(\mathbf{x}', x_n; E) = -\delta(\mathbf{x} - \mathbf{y}).$$

The Fourier transform from  $E$  (energy) to  $t$  (time) turns this into

$$(\Delta + i\partial_t + x_n)\tilde{\Phi}(\mathbf{x}', x_n; t) = -\delta(\mathbf{x} - \mathbf{y})\delta(t) \quad (9)$$

which is the fundamental solution for the time-dependent Schrödinger equation in a linear potential. We may solve this exactly by performing a Fourier transform in space, from coordinates  $(\mathbf{x}', x_n)$  to wavevector  $(\mathbf{k}', K)$ ,

$$\left[ -|\mathbf{k}'|^2 - K^2 + i(\partial_t + \partial_K) \right] \hat{\Phi}(\mathbf{k}', K; t) = -(2\pi)^{-n} e^{-i\mathbf{k}' \cdot \mathbf{y}} \delta(t).$$

The only derivatives are an advection term causing constant unit speed drift in wavevector in the positive  $x_n$  direction, so we shift to a frame moving in wavevector, substituting  $\kappa = K - t$  (in physics this is called a gauge change [8, App. A]). To change from coordinates  $(\mathbf{k}', K; t)$  to  $(\mathbf{k}', \kappa; t)$  we then need

$$\left. \frac{\partial}{\partial t} \right|_\kappa = \left. \frac{\partial}{\partial t} \right|_K + \frac{\partial}{\partial K}.$$

This gives the simple first-order ODE in time at each wavevector  $(\mathbf{k}', \kappa) \in \mathbb{R}^n$ ,

$$\left[ -|\mathbf{k}'|^2 - (\kappa + t)^2 + i\partial_t \right] \hat{\Phi}(\mathbf{k}', \kappa; t) = -(2\pi)^{-n} e^{-i(\mathbf{k}' \cdot \mathbf{y}' + \kappa y_n)} \delta(t).$$

For each  $(\mathbf{k}', \kappa) \in \mathbb{R}^n$  we seek a causal solution with  $\hat{\Phi}(\mathbf{k}', \kappa; t) = 0$  for all  $t < 0$ . The right-hand side is an impulsive excitation at  $t = 0$  which gives the ODE solution

$$\hat{\Phi}(\mathbf{k}', \kappa; t) = \frac{ie^{-i(\mathbf{k}' \cdot \mathbf{y}' + \kappa y_n)}}{(2\pi)^n} \exp i \left[ -|\mathbf{k}'|^2 t - \kappa^2 t - \kappa t^2 - \frac{1}{3} t^3 \right], \quad t > 0.$$

Changing back to the original wavevector coordinates via  $\kappa = K - t$  gives

$$\hat{\Phi}(\mathbf{k}', K; t) = \frac{ie^{i(y_n t - \frac{1}{3} t^3)}}{(2\pi)^n} e^{-i\mathbf{k} \cdot \mathbf{y}} \exp i \left[ -|\mathbf{k}'|^2 t - K^2 t + K t^2 \right], \quad t > 0.$$

The final exponential is an (imaginary) gaussian in Fourier space, whose inverse spatial Fourier transform is known exactly. The middle exponential term causes a real space translation by  $\mathbf{y}$ . This gives after simplification,

$$\tilde{\Phi}(\mathbf{x}, \mathbf{y}; t) = \frac{i}{(4\pi i t)^{n/2}} \exp i \left[ \frac{|\mathbf{x} - \mathbf{y}|^2}{4t} + \frac{x_n + y_n}{2} t - \frac{1}{12} t^3 \right], \quad t > 0. \quad (10)$$

This is the fundamental solution to the time-dependent Schrödinger equation (9). An inverse Fourier transform in time returns to the frequency-domain, giving the desired (8).  $\square$

**Remark 2** (plain Helmholtz equation). *Applying the above technique to the constant-wavenumber Helmholtz equation  $(\Delta + E)u = 0$  gives the fundamental solution representation*

$$\Phi(\mathbf{x}, \mathbf{y}) = \frac{i}{(4\pi i)^{n/2}} \int_0^\infty \frac{1}{t^{n/2}} \exp i \left[ \frac{|\mathbf{x} - \mathbf{y}|^2}{4t} + Et \right] dt,$$

which is the same as (8) absent two terms. In the case  $n = 2$ , by changing variable to  $s = (2i\sqrt{E}/r)t$ , where  $r = |\mathbf{x} - \mathbf{y}|$ , we see that the above is the little-known Schlöfli integral representation [37, (4) Sec. 6.21] for the radiative fundamental solution  $(i/4)H_0^{(1)}(\sqrt{E}r)$ , where  $H_0^{(1)}$  is the outgoing Hankel function of order zero.

**Remark 3.** *We leave for future work a proof that the causal fundamental solution (8) satisfies our radiation condition in Definition 1, although physical intuition, the Helmholtz case, and numerical evidence strongly suggest that this is the case. A proof seems to demand stationary phase estimates beyond the scope of this work. A rigorous existence proof for the BVP (1)–(3) would follow, in an analogous fashion to [22, Thm. 3.9].*

The importance of (8) is that quadrature of this integral will provide us with an accurate numerical algorithm to evaluate the fundamental solution for  $n = 2$  (Section 4).

Finally we recall a property of  $\Phi$  special to  $n = 2$ . Since the PDE has coefficients which vary as analytic functions of  $x_1$  and  $x_2$ , the fundamental solution must have the form [38, Ch. 5]

$$\Phi(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y}) \frac{1}{2\pi} \log \frac{1}{|\mathbf{x} - \mathbf{y}|} + B(\mathbf{x}, \mathbf{y}), \quad (11)$$

where  $A$  and  $B$  are analytic in both coordinates of both variables, and  $A(\mathbf{x}, \mathbf{x}) = 1$  for all  $\mathbf{x} \in \mathbb{R}^2$ . Thus, as with the Laplace and Helmholtz equations, there is a (positive sign) logarithmic singularity at the source point.

### 2.1. Connection to ray dynamics, propagating and forbidden regions

In Fig. 2 we plot the fundamental solution, showing the different behaviors resulting by varying the height of the source location  $\mathbf{y}$  at fixed energy  $E$ . In panel (a) the radiation from the source point is visible, as is interference between upwards and downwards propagating waves. In panel (b) the source is closer to the turning height  $x_2 = -E$ , and ray trajectories have been superimposed showing the connection to classical dynamics. We now review this connection (see e.g. [39, Sec. 9-10], [40, Sec. 4.5], [6, Sec. 5.1] [41]). Consider the general variable-coefficient Helmholtz equation

$$(\Delta + k(\mathbf{x})^2)u = 0.$$

When  $k$  is locally large, inserting Keller's traveling wave ansatz  $u(\mathbf{x}) = a(\mathbf{x})e^{i\phi(\mathbf{x})}$  into the PDE gives to leading order the eikonal equation  $|\nabla\phi| = k(\mathbf{x})$ , whose characteristics are rays given by evolving Hamilton's equations (here a dot indicates a time derivative),

$$\dot{\mathbf{x}} = \nabla_{\mathbf{p}}H, \quad \dot{\mathbf{p}} = -\nabla_{\mathbf{x}}H, \quad (12)$$

with the Hamiltonian  $H(\mathbf{x}, \mathbf{p}) = |\mathbf{p}|^2 + V(\mathbf{x})$  and potential  $V(\mathbf{x}) = -k(\mathbf{x})^2 + E$ , with (conserved) total energy  $H = E$ , where  $E$  is any constant. (Here the kinetic energy term corresponds to a particle of mass  $\frac{1}{2}$ .) Another way to express this is via quantization, or "quantum-classical correspondence", which associates the operator  $i\nabla$  with the momentum variable  $\mathbf{p}$ . This rigorous connection is the topic of semiclassical analysis [42].

Returning to our case of stratified  $k$ , and the constant  $E$ , given by (3), then  $V(\mathbf{x}) = -x_2$ , we see that rays evolve under a constant "gravitational" force field  $-\nabla V(\mathbf{x}) = (0, 1)$  in the vertical direction, i.e. Hamilton's equations are  $\dot{\mathbf{x}} = 2\mathbf{p}$  and  $\dot{\mathbf{p}} = (0, 1)$ . To model the fundamental solution  $\Phi(\cdot, \mathbf{y})$ , rays are launched from the source  $\mathbf{y}$ , with initial momentum  $\boldsymbol{\rho} = (\rho_1, \rho_2)$ , hence have the Galilean solution

$$x_1(t) = y_1 + 2\rho_1 t, \quad x_2(t) = y_2 + 2\rho_2 t + t^2. \quad (13)$$

Fig. 2 suggests that such rays predict the wavefronts and caustics of  $\Phi$ , and that  $\Phi$  is small in the "classically forbidden" region, which we call region F, defined in the following.

**Proposition 1.** *Rays obeying (12) with Hamiltonian  $H(\mathbf{x}, \mathbf{p}) = |\mathbf{p}|^2 - x_2$  launched from  $\mathbf{y}$  with total energy  $E$  cannot reach the forbidden region F, which is defined by  $\mathbf{x} = (x_1, x_2)$  such that*

$$\frac{|\mathbf{x} - \mathbf{y}|}{2} > \frac{x_2 + y_2}{2} + E, \quad (14)$$

whose boundary is the parabola with focus  $\mathbf{y}$  and directrix  $x_2 = -y_2 - 2E$ . Rays can reach any point in the complement of the region, which we will label region A, for "classically allowed".

We provide a proof, simplifying that of Bracher et al. [8], that introduces the concept of *travel time*, crucial to the later numerical evaluation.

*Proof.* We substitute the formulae for  $\rho_1$  and  $\rho_2$  from (13) into the expression  $\rho_1^2 + \rho_2^2 = E + y_2$  expressing that the initial total energy  $H(\mathbf{y}, \boldsymbol{\rho}) = E$ , to get the quadratic equation in  $t^2$ ,

$$\frac{t^4}{4} - bt^2 + a = 0 \quad (15)$$

where for later simplicity we define

$$a := \frac{|\mathbf{x} - \mathbf{y}|^2}{4}, \quad b := \frac{x_2 + y_2}{2} + E. \quad (16)$$

The positive solutions to (15) give possible ray travel times from  $\mathbf{y}$  to  $\mathbf{x}$  at fixed  $E$ , being

$$t_{\pm} = +\sqrt{2(b \pm \sqrt{b^2 - a})}. \quad (17)$$

No real solutions are possible precisely when  $\sqrt{a} > b$ , which gives (14). The boundary, written  $2\sqrt{a} = 2b$ , states that the distance from  $\mathbf{y}$  to  $\mathbf{x}$  equals the distance from  $\mathbf{x}$  to the directrix line  $x_2 = -y_2 - 2E$ , defining a parabola.  $\square$

At the parabolic boundary the two travel times coalesce, i.e.  $t_- = t_+$ , causing a caustic (singularity in density) for the rays, which manifests itself as large amplitudes in the fundamental solution; see Fig. 2(a)–(b). We show in Fig. 2(c) a case where the source itself lies in the forbidden region. Here there are no classical rays and the wave leakage into the propagating region is exponentially small, occurring only in a single upwards direction.

Finally we emphasize that time evolution appears in two different settings in this section: in the time-dependent Schrödinger equation to give  $t$  in the integral (8), and the time variable  $t$  in the classical dynamics. We have chosen the dimensionless units (i.e. particle mass  $\frac{1}{2}$ ) so that they correspond.

### 3. Conversion to a boundary integral equation, and its numerical solution

We will reformulate the exterior Dirichlet BVP (1)–(3) as a Fredholm second-kind integral equation on  $\partial\Omega$ . Since it provides us a useful numerical test case, we also do the same for the interior BVP. Recall that by standard elliptic PDE theory, given a compact domain  $\Omega$ , the interior Dirichlet BVP has a unique solution for all  $E$  except at a countable set (the Dirichlet eigenvalues of the operator  $-\Delta - x_2$ ) that accumulates only at infinity [43, Thms. 4.10, 4.12].

Given the fundamental solution  $\Phi(\mathbf{x}, \mathbf{y})$ , and a “density” function  $\tau$  on the boundary curve  $\partial\Omega$ , we define the standard single- and double-layer potential representations,

$$(\mathcal{S}\tau)(\mathbf{x}) := \int_{\partial\Omega} \Phi(\mathbf{x}, \mathbf{y})\tau(\mathbf{y}) ds_{\mathbf{y}}, \quad (\mathcal{D}\tau)(\mathbf{x}) := \int_{\partial\Omega} \frac{\partial\Phi(\mathbf{x}, \mathbf{y})}{\partial\mathbf{n}_{\mathbf{y}}} \tau(\mathbf{y}) ds_{\mathbf{y}}, \quad (18)$$

where  $\mathbf{n}(\mathbf{y})$  is the outward-pointing unit normal vector at the point  $\mathbf{y} \in \partial\Omega$ , and  $ds$  the usual arc length element. One may interpret  $\mathbf{y}$  as a source point and  $\mathbf{x}$  as a target. Since limits of such potentials on the curve itself may depend on from which side it is approached, we define

$$v^{\pm}(\mathbf{x}) := \lim_{h \rightarrow 0^+} v(\mathbf{x} \pm h\mathbf{n}(\mathbf{x})).$$

Letting  $S : C(\partial\Omega) \rightarrow C(\partial\Omega)$  be the boundary integral operator with kernel  $\Phi(\mathbf{x}, \mathbf{y})$ , and  $D : C(\partial\Omega) \rightarrow C(\partial\Omega)$  be the boundary integral operator with kernel  $\partial\Phi(\mathbf{x}, \mathbf{y})/\partial\mathbf{n}(\mathbf{y})$  taken in the principal value sense, we have jump relations,

$$(\mathcal{D}\tau)^{\pm}(\mathbf{x}) = (D\tau \pm \tfrac{1}{2}\tau)(\mathbf{x}), \quad (19)$$

$$(\mathcal{S}\tau)^{\pm}(\mathbf{x}) = (S\tau)(\mathbf{x}), \quad (20)$$

which are identical to the Laplace and Helmholtz cases [22, Thm. 3.1 and p.66]. For the proof we need the variable-coefficient elliptic PDE case [43, Thm. 6.11 and (7.5)].

The indirect BIE is constructed by making the “combined field integral equation” (CFIE) ansatz

$$u = (\mathcal{D} - i\eta\mathcal{S})\tau \quad (21)$$

and substituting this into the boundary condition (2), using the exterior jump relations to get the BIE for the unknown density  $\tau$ ,

$$(\tfrac{1}{2}I + D - i\eta S)\tau = f \quad \text{exterior BIE}, \quad (22)$$

where  $I$  is the identity. This mixture of double- and single-layer prevents a spurious resonance problem (for  $\eta = 0$  the operator would be singular at interior Neumann eigenvalues), making the BIE a robust method for the BVP. The choice of constant  $\eta$  is not crucial but is commonly scaled with the wavenumber [27]; our wavenumber varies in space, and we choose at typical value  $\eta = \sqrt{E}$ . Note that the correct sign of  $\eta$  is crucial for rapid convergence of iterative solvers at high frequency.

For the interior BVP, the CFIE is not (usually) needed, so we set  $\eta = 0$  and get

$$(-\tfrac{1}{2}I + D)\tau = f \quad \text{interior BIE}. \quad (23)$$

Note that the operator  $S$  is compact, and when  $\partial\Omega$  is smooth the operator  $D$  is compact, making the above BIEs of Fredholm second kind. This has the well-known advantages over first-kind BIEs of stability under discretization, and a benign spectrum leading to rapid convergence for the iterative solution of the resulting linear system.

#### 3.1. Numerical solution: Nyström method and quadrature

We first parametrize the smooth closed curve  $\partial\Omega$  by a  $2\pi$ -periodic function  $\mathbf{z} : [0, 2\pi) \rightarrow \mathbb{R}^2$  such that  $\mathbf{z}(t) \in \partial\Omega$  and  $|\mathbf{z}'(t)| \neq 0$ , for all  $t \in \mathbb{R}$ . Changing variable to the parameter  $t$  turns (22) into an integral equation on the periodic interval  $[0, 2\pi)$ ,

$$\tfrac{1}{2}\tau(t) + \int_0^{2\pi} \left( \frac{\partial\Phi(\mathbf{z}(t), \mathbf{z}(s))}{\partial\mathbf{n}_{\mathbf{z}(s)}} - i\eta\Phi(\mathbf{z}(t), \mathbf{z}(s)) \right) |\mathbf{z}'(s)| \tau(s) ds = f(t), \quad \forall t \in [0, 2\pi) \quad (24)$$



The reparametrization of (23) is similar. We can write both of these integral equations in the standard form

$$\tau(t) + \int_0^{2\pi} K(t, s)\tau(s)ds = g(t), \quad \forall t \in [0, 2\pi) \quad (25)$$

In the exterior case, we see from the presence of  $\Phi$  and from (11) that  $K$  has a logarithmically singular kernel, i.e.  $K(s, t) \sim \log|s - t|$ ; in the interior case the kernel of  $K$  is continuous at the diagonal but has a weaker singularity of the form  $|s - t|^2 \log|s - t|$ , as with the Helmholtz equation [22, Sec. 3.5]. To achieve high-order convergence in either case when the data  $g$  is smooth we will need to use a quadrature scheme accurate for kernels containing a periodized log singularity of the form

$$K(t, s) = K_1(t, s) \log\left(4 \sin^2 \frac{s-t}{2}\right) + K_2(t, s) \quad (26)$$

where  $K_1$  and  $K_2$  are smooth and  $2\pi$ -periodic in both of their arguments.

We apply the Nyström method [44, Sec. 12.3] to approximate the solution of (25) by that of a linear system, based upon an underlying quadrature rule. For this we use periodic trapezoid rule quadrature,

$$\int_0^{2\pi} \phi(t)dt \approx \frac{2\pi}{N} \sum_{j=1}^N \phi(s_j), \quad \text{where } s_j = 2\pi j/N \quad (27)$$

whose approximation error for a  $2\pi$ -periodic  $\phi \in C^\infty(\mathbb{R})$  is super-algebraic, i.e.  $O(N^{-m})$  for each  $m > 0$  [44, Cor. 9.27]. The first step in the Nyström method is to enforce (25) only at the nodes  $\{s_i\}$ , giving

$$\tau(s_i) + \int_0^{2\pi} K(s_i, s)\tau(s)ds = g(s_i), \quad \forall i = 1, \dots, N \quad (28)$$

Were  $K$  to possess a smooth kernel (i.e.  $K_1 \equiv 0$ ), superalgebraic convergence would be achieved by applying (27) to the above integral, to give the square  $N$ -by- $N$  linear system,

$$\tau_i + \sum_{j=1}^N A_{ij}\tau_j = g_i, \quad \forall i = 1, \dots, N \quad (29)$$

with elements of the matrix given by

$$A_{ij} = \frac{2\pi}{N} K(s_i, s_j), \quad (30)$$

and where  $\tau_j$  approximates  $\tau(s_j)$  and the right-hand side vector has elements  $g_j = g(s_j)$ .

However, for general singular kernels of the form (26), the formula (30) fails to be accurate, and diagonal entries would be infinite. Yet it is still possible to design a set of quadrature nodes to approximate the integral in (28) to high accuracy for kernels of the form (26). This is done by replacing a few of the trapezoid nodes  $s_j$  near the singularity  $s_i$  by a new set of auxiliary nodes and weights; we choose 16th-order Alpert end-correction nodes [45], of which 30 are required (15 either side of the singularity). The auxiliary node nearest the target point is at a distance of around  $10^{-3}\delta$  from this target point, where  $\delta \approx (2\pi/N)|\mathbf{z}'(s_i)|$  is the local underlying node spacing. The values of  $\tau$  at these auxiliary nodes is related to the neighboring few elements of the vector  $\{\tau_j\}_{j=1}^N$  using local Lagrange interpolation. The net effect is that the matrix  $A$  takes the form (30) away from the diagonal, but with corrected entries near the diagonal. The full formulae are presented in [28, Sec. 4]. This gives for kernels of the form (26) a high-order convergence of the error between  $\tau_j$  and the true solution samples  $\tau(s_j)$  of  $O(N^{-16} \log N)$ , for either the exterior or interior BIEs of interest. For the convergence theory see [45, Cor. 3.8] for the end-correction scheme, and Kress [46, Ch. 12].

Once the linear system (29) has been solved, the vector  $\tau = \{\tau_j\}_{j=1}^N$  may be used to reconstruct the scattered potential at any target location sufficiently far from  $\partial\Omega$ , by substituting the same trapezoid rule into the integrals (18) in the representation (21), to get

$$u(\mathbf{x}) = \sum_{j=1}^N \left( \frac{\partial\Phi(\mathbf{x}, \mathbf{z}(s_j))}{\partial\mathbf{n}_{\mathbf{z}(s_j)}} - i\eta\Phi(\mathbf{x}, \mathbf{z}(s_j)) \right) |\mathbf{z}'(s_j)| \tau_j \quad (31)$$

A rule of thumb is that this quadrature rule is accurate for all points at least  $5\delta$  from the boundary [47, Remark 6]. As before, for the interior case we set  $\eta = 0$ .

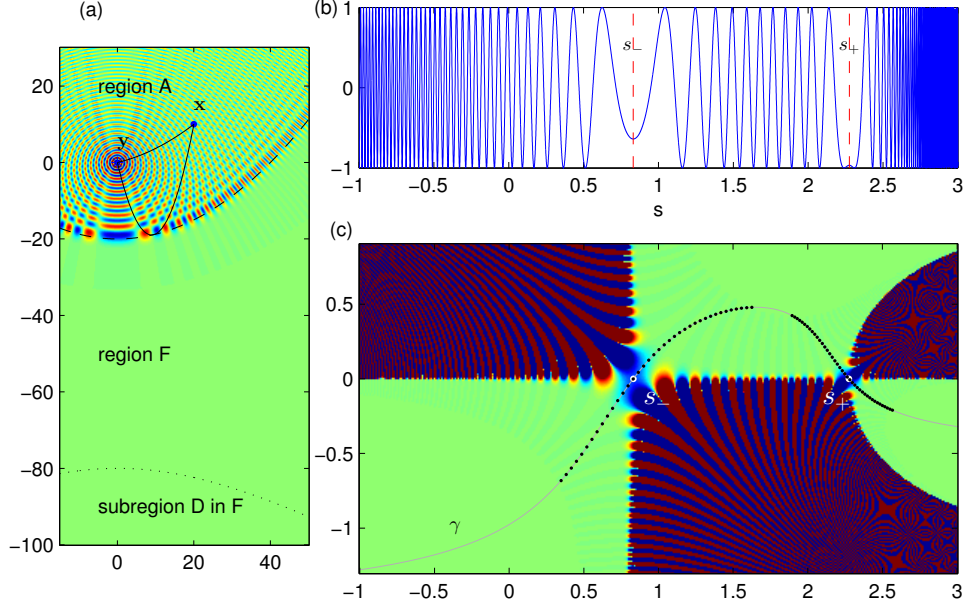


Figure 3: Contour integration for the fundamental solution at  $E = 20$ ,  $\mathbf{y} = \mathbf{0}$ . (a)  $\text{Re } \Phi(\mathbf{x}, \mathbf{y})$  in the physical domain of  $\mathbf{x}$ , showing the two ray paths to reach  $\mathbf{x} = (20, 10)$ , and the classically allowed (A), forbidden (F), and deep forbidden (D) regions. (b) integrand of (34) on the real  $s$  axis, with the two stationary phase points  $s_{\pm} = \log t_{\pm}$ . (c) real part of the same integrand in the complex  $s$  plane, saddle points (white dots), and the 79 quadrature nodes used lying on the contour  $\gamma$ . In (a) and (c) the color scale is blue (negative) through green (zero) to red (positive); in (c) the color range covers  $[-1, 1]$ .

#### 4. Evaluation of the fundamental solution

Filling the Nyström matrix  $A$  of the previous section, and evaluating the solution  $u$  via (31), both demand a large number of evaluations of  $\Phi(\mathbf{x}, \mathbf{y})$ , from source points  $\mathbf{y}$  that are either periodic trapezoid nodes  $\mathbf{z}(s_j)$  or auxiliary nodes. When filling  $A$  the target points  $\mathbf{x}$  are also the nodes  $\mathbf{z}(s_i)$ , thus for a small number of cases ( $O(N)$  of them), the distance  $|\mathbf{x} - \mathbf{y}|$  will be very small (e.g.  $10^{-3}\delta$ ).

As promised, we base our evaluation of the fundamental solution on the  $n = 2$  dimensional case of (8),

$$\Phi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \int_0^\infty \frac{1}{t} \exp i \left[ \frac{|\mathbf{x} - \mathbf{y}|^2}{4t} + \left( \frac{x_n + y_n}{2} + E \right) t - \frac{1}{12} t^3 \right] ds = \frac{1}{4\pi} \int_0^\infty \frac{1}{t} e^{i\psi(t)} dt \quad (32)$$

where, recalling (16), the *phase function*  $\psi(t) = \psi_{a,b}(t)$  is defined by

$$\psi(t) := \frac{a}{t} + bt - \frac{1}{12} t^3. \quad (33)$$

To remove the pole at the origin, and place small and large  $t$  on an equal footing, we change variable via  $t = e^s$  to get

$$\Phi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \int_{-\infty}^\infty \exp i\psi(e^s) ds. \quad (34)$$

This integrand is shown in Fig. 3(b), for  $E = 20$  and the source  $\mathbf{y}$  and target  $\mathbf{x}$  shown in Fig. 3(a). It is clearly highly oscillatory—and it becomes more so with increasing  $E$ —thus accurate integration along the real  $s$  axis would be prohibitively expensive. However,  $\phi(e^s)$ , and hence the integrand, is analytic in the entire complex  $s$  plane. We thus use numerical saddle point integration [14, Sec. 5.5] [48] (related to, but simpler than, “numerical steepest descent” [49]), along a contour passing through the stationary phase (saddle) points and asymptotically tending to the correct regions of the plane. We have the following by direct differentiation of (33).

**Proposition 2.** *Given a source  $\mathbf{y}$ , target  $\mathbf{x}$ , and energy  $E$ , the stationary phase points, that is, the solutions to  $\psi'(t) = 0$ , are precisely the classical ray travel times  $t_{\pm}$  already given by (16)-(17).*

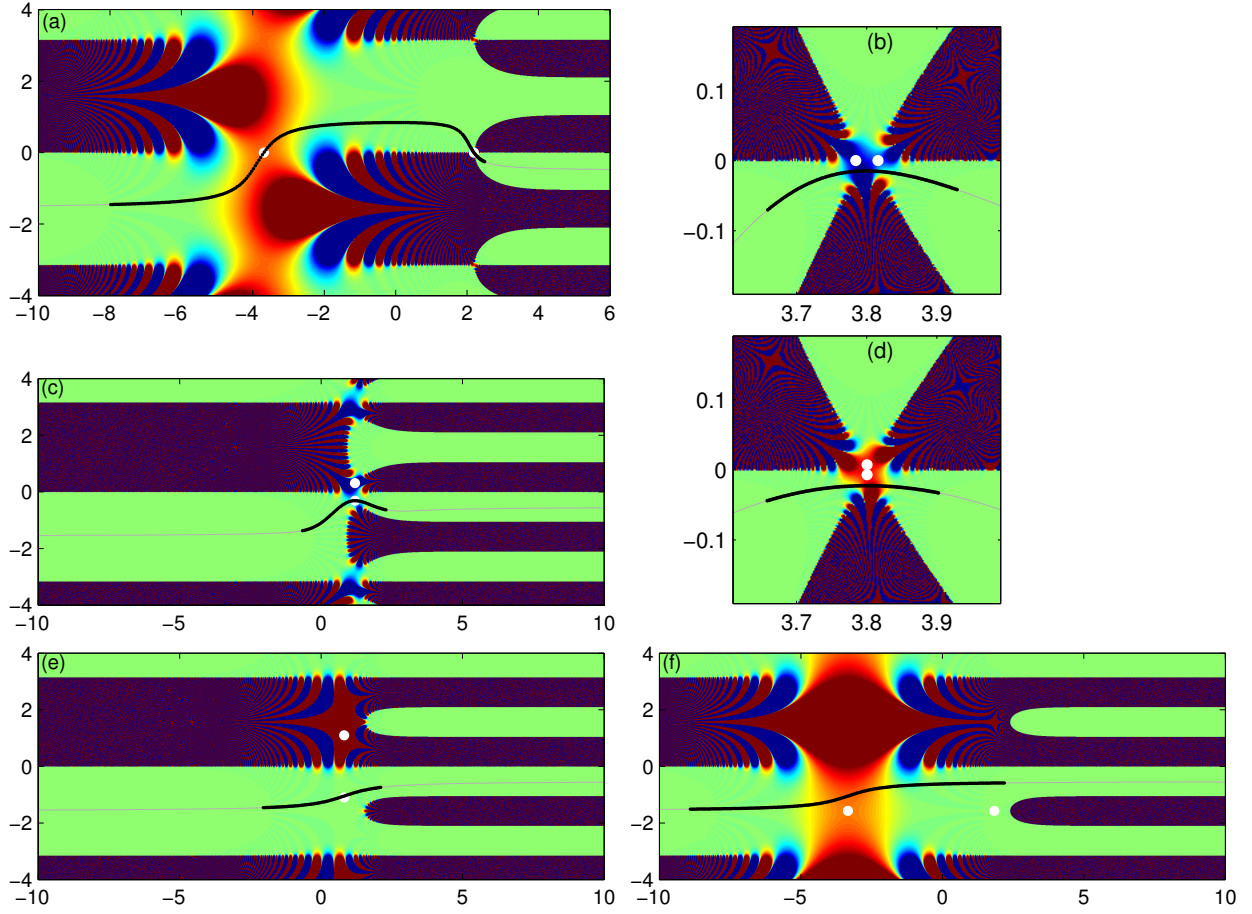


Figure 4: Real part of integrands plotted in the complex  $s$  plane, for source  $\mathbf{y} = \mathbf{0}$ , with saddle points (white dots) and numerical integration contours (grey) and nodes (black). (a) Region A (allowed) but source close to target,  $E = 20$ ,  $\mathbf{x} = (0.1, 0.2)$ . (c) Region F (forbidden),  $E = 10$ ,  $\mathbf{x} = (1, -11)$ . (b) and (d) Zoom in on coalescing saddle points: at  $E = 10^3$ , with  $\mathbf{x} = (2E - 1, 0)$  in (b) (just allowed), and  $\mathbf{x} = (2E + 0.2, 0)$  in (d) (just forbidden). (e) Region D (deep forbidden),  $E = 1$ ,  $\mathbf{x} = (1, -5)$ . (f) Region D but source close to target,  $E = -10$ ,  $\mathbf{x} = (0.1, 0.2)$ .

This connection between waves and rays is key to our efficient numerical evaluation of the integral (34).

**Remark 4.** *There is a beautiful and deep physical reason lying behind Prop. 2, i.e.  $\psi'(t_{\pm}) = 0$ . The phase function (term in square brackets) in the time-dependent Schrödinger propagator (10) is the classical action  $S(\mathbf{x}, \mathbf{y}; t)$ , defined as the time integral over  $[0, t]$  of the Lagrangian along the unique classical path from  $\mathbf{y}$  to  $\mathbf{x}$  taking precisely time  $t$  [8, Sec. 2] [41, Ch. 10]. (Note that for a general potential function  $V(\mathbf{x})$ , this is only approximately true in the semi-classical or high-frequency limit; its exactness here reflects exact formulae for the propagation of the Gaussian when the potential is at most quadratic in the coordinates [13].)*

*Inserting this into the last step in the proof of Lemma 1, we see that  $\Phi(\mathbf{x}, \mathbf{y}) = \int_0^\infty \exp i[S(\mathbf{x}, \mathbf{y}; t) + Et] dt$ , thus the phase function (33) is  $\psi(t) = S(\mathbf{x}, \mathbf{y}; t) + Et$ . A less well-known result from classical mechanics is  $\partial S(\mathbf{x}, \mathbf{y}; t)/\partial t|_{\mathbf{x}, \mathbf{y}} = -E_{\mathbf{x}, \mathbf{y}}(t)$ , where  $E_{\mathbf{x}, \mathbf{y}}(t)$  is the energy required to complete the path in time  $t$ . [41, Ex. 10.4(c)]. Thus  $\psi'(t) = 0$  precisely when  $E_{\mathbf{x}, \mathbf{y}}(t) = E$ , that is, at the travel times for a ray at the particular energy  $E$  to pass from  $\mathbf{y}$  to  $\mathbf{x}$ .*

An example contour passing through the (real-valued) saddle points and ending in the correct regions of the plane is shown in Fig. 3(c). On such a contour the integral may be approximated to exponential accuracy using the trapezoid rule [15] (with respect to the variable parametrizing the contour), and the sum may be truncated once values are sufficiently small.

#### 4.1. Choice of saddle point contour

Since the integrand in (34) is entire, mathematically the choice of contour is irrelevant as long as its ends connect  $-\infty$  to  $+\infty$ . However, for practical numerical evaluation the contour choice is crucial. Observe in Fig. 3(c) that the integrand is exponentially small in some regions, exponentially large in others, and that the borders between them are quite well defined. One may deform the limits of the contour to lie below the real axis, as long as one stays within the exponentially small regions adjoining the real axis (lower-left and lower-right in Fig. 3(c)). It must connect these limits, but to prevent catastrophic cancellation it must avoid large regions, passing between small regions only via saddle points, and passing through these saddle points at an angle not too far from the steepest descent direction. In addition, an analytic contour shape is desirable, since the trapezoid rule is then exponentially convergent. See Fig. 3(c) and Fig. 4 for examples.

The task remains to choose, for any parameters  $a$  and  $b$ , a good contour, and rules for choosing the trapezoid node spacing and truncation intervals. Our rules will depend on the existence and types of classical rays. Recall the definition that the set  $\mathbf{x}$ ,  $\mathbf{y}$  and  $E$  is classically allowed (region A) if there is one or two rays connecting  $\mathbf{y}$  to  $\mathbf{x}$  at energy  $E$  in (real-valued) time, otherwise forbidden (region F).

##### 4.1.1. Classically allowed (region A): $b^2 \geq a$

In this case, as in Fig. 3, there are two real saddle points, with steepest descent angles  $\pi/4$  for  $s_- = \log t_-$  (the root with smaller real part), and  $-\pi/4$  for  $s_+ = \log t_+$ . We parametrize contours by their real part  $\alpha \in \mathbb{R}$ , thus

$$s = \gamma(\alpha) := \alpha + ig(\alpha), \quad \text{hence} \quad \gamma'(\alpha) = 1 + ig'(\alpha),$$

where the function  $g : \mathbb{R} \rightarrow \mathbb{R}$  depends on the usual parameters  $a$  and  $b$  (16). The following analytic function  $g$  makes the contour pass through the two saddle points at angles not too far from  $\pm\pi/4$ ,

$$g(\alpha) = \left[ \left( \frac{1}{\pi} + \frac{1}{2} \right) \tan^{-1}(2(\alpha - \operatorname{Re} s_- + c_-)) - \left( \frac{\pi}{4} - \frac{1}{2} \right) \right] \cdot \left[ \left( \frac{1}{\pi} + \frac{1}{6} \right) \tan^{-1}(-4(\alpha - \operatorname{Re} s_+ - c_+)) - \left( \frac{\pi}{12} - \frac{1}{2} \right) \right] \quad (35)$$

with the constants  $c_- := \frac{1}{2} \tan(\frac{\pi^2 - 2\pi}{4 + 2\pi})$  and  $c_+ := \frac{1}{4} \tan(\frac{\pi^2 - 6\pi}{12 + 2\pi})$ . We do not claim it is optimal, but it serves our purpose well. Examples from this family are shown in Fig. 3(c) and Fig. 4(a).

The leftward limit  $\lim_{\alpha \rightarrow -\infty} g(\alpha) = -\pi/2$  is designed to lie in the middle of the exponentially-small region to the left. This region has height  $\pi$  due to the  $2\pi$  vertically periodic nature of the function  $e^{-s}$  which dominates as  $\operatorname{Re} s$  becomes highly negative. To the right the period becomes three times smaller, since  $e^{3s}$  is dominant, thus we chose  $\lim_{\alpha \rightarrow \infty} g(\alpha) = -\pi/6$ . Note that it is essential to enter and exit through the correct periodic images on the left and right sides.

When the saddle points coalesce ( $t_- = t_+$  at the classical turning point, or boundary of A and F), the angles through the saddle points become flatter, as is needed to traverse smoothly through the small region; see the zoom Fig. 4(b). However, when saddles are close to coalescing at high  $E$ , it is advantageous for accuracy to shift the contour down enough to avoid being close to the rapid oscillations on the real axis, whilst keeping the integrand not too large. Hence, when  $|s_+ - s_-| < 0.1$  we add the constant

$$c_{\text{shift}} := -i \min \left[ \frac{0.7}{\sqrt{E}}, 0.1 \right] \quad (36)$$

to  $\gamma$ . The resulting shift is visible in the figure.

##### 4.1.2. Classically forbidden (region F): $b^2 < a$

Things get simpler when no real rays are possible: the saddle points  $s_{\pm}$  split away from the real  $s$  axis, and only the one with negative imaginary part is relevant. Let us call this point  $s_0$ . There are a couple of regimes to consider; see Fig. 4(c)–(f). We use the following contour when  $\operatorname{Im} s_0 > -\pi/3$ ,

$$g(\alpha) = \operatorname{Im} s_0 + \left( \frac{\tan^{-1}(\alpha - \operatorname{Re} s_0) - \pi}{3} - \operatorname{Im} s_0 \right) (1 - e^{-(\alpha - \operatorname{Re} s_0)^2}).$$

This has the same limits as (35), is designed to pass through  $s_0$  horizontally (i.e.  $g'(\operatorname{Re} s_0) = 0$ ), and is shown in Fig. 4(c). The need for horizontal passage is to stay below the real axis when saddles are close to coalescing at high  $E$ . As above, we also apply the shift (36) when saddle points are close. This is shown in the zoom Fig. 4(d).

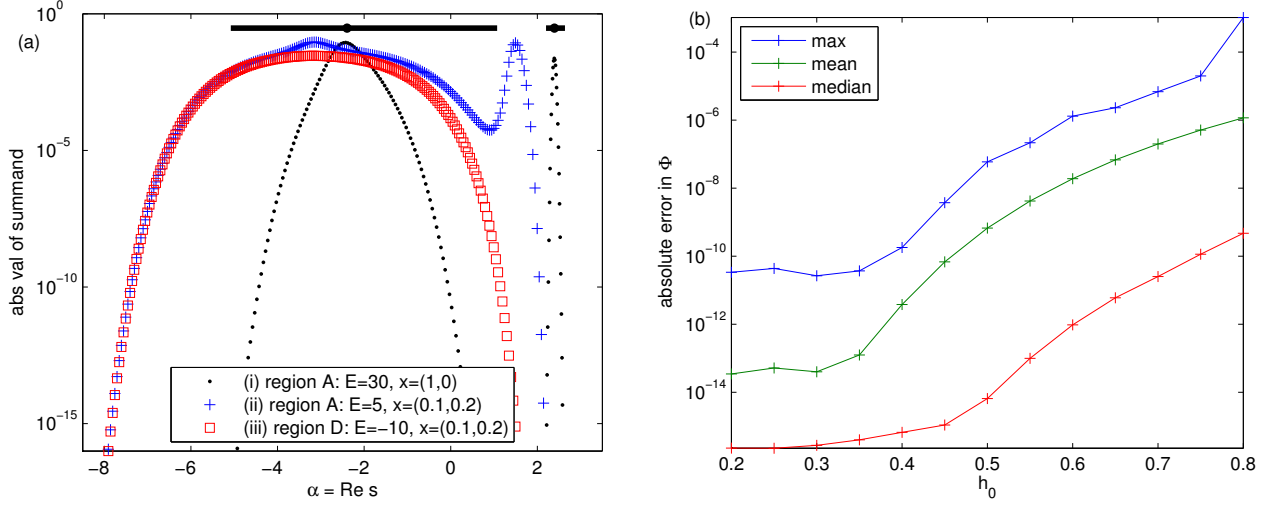


Figure 5: (a) Magnitude of summand in (38) along the parametrized contour, showing three types of behavior. For case (i) the intervals  $I_1$  and  $I_2$  containing the saddle points (large dots) are shown at the top. (b) Convergence of absolute error in  $\Phi$ , with respect to the quadrature spacing  $h_0$ , also scaling  $h_{\max} = 0.13 h_0$  and  $n_{\min} = 15/h_0$ . The source is  $\mathbf{y} = \mathbf{0}$ , and targets are a set of  $10^4$  points randomly distributed uniformly in angle and uniformly in the logarithm of distance from the origin,  $|\mathbf{x}| \in [10^{-4}, 10^4]$ . For each target the set of  $E$  tested is  $[-100, -30, -10, -3, -1, 1, 3, 10, 30, 100, 300, 10^3, 3 \times 10^3, 10^4]$ . The maximum, mean, and median error is taken over the  $1.5 \times 10^5$  evaluations.

When  $\text{Im } s_- < -\pi/3$ , we are deep into the forbidden region (thus we call the region  $D \subset F$ ). It lies below the hyperbola  $b = -\sqrt{a}/2$  in the  $\mathbf{x}$  plane, as shown in Fig. 3(a). In region D we use the simple contour

$$g(\alpha) = \frac{\tan^{-1}(\alpha - \text{Re } s_0) - \pi}{3}.$$

This lies above all saddle points, has the same limits as (35), and is shown in Fig. 4(e).

When  $b < -\sqrt{a}$ , as occurs in region D with negative  $E$  and close source-target distances, the saddle points finally merge again onto the line  $\text{Im } s = -i\pi/2$ . In this case we take  $s_0$  to be the point with more negative real part, and use the above contour. This is shown in Fig. 4(f).

#### 4.2. Truncation of the integration domain

With contour shapes now defined for all cases of  $a$  and  $b$ , we need rules to truncate the integral to a finite domain  $I \subset \mathbb{R}$ , that is,

$$\Phi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \exp i\psi(e^s) ds = \frac{1}{4\pi} \int_{-\infty}^{\infty} \exp i\psi(e^{\gamma(\alpha)}) \gamma'(\alpha) d\alpha \approx \frac{1}{4\pi} \int_I \exp i\psi(e^{\gamma(\alpha)}) \gamma'(\alpha) d\alpha. \quad (37)$$

For efficiency, we wish  $I$  to enclose only the parts of  $\mathbb{R}$  where the integrand is significant, which we define as exceeding a convergence parameter  $\varepsilon$ , which we set to  $10^{-14}$ . We exploit the fact that, along the contour, the integrand decays exponentially away from saddle points.

There are three types of behavior: (i)  $I$  comprises two intervals  $I_1$  and  $I_2$  that may be integrated independently, (ii) there are two saddle points but the integrand does not die to  $\varepsilon$  between them, so it must be handled as a single integration interval, and (iii) there is one saddle point hence only a single “bump” and a single interval. For case (i), for high  $E$  the size of the intervals can be much smaller than their separation, so integrating them separately is crucial. All three cases are shown in Fig. 5(a). In region A, (i) and (ii) may occur; in regions F and D only (iii) occurs.

The recipe for regions F and D, with one saddle  $s_0$ , is to initialize distances  $d_1 = d_2 = |\text{Re } s_0|/2$  which define an interval  $[\text{Re } s_0 - d_1, \text{Re } s_0 + d_2]$ . If  $|\psi(e^{\gamma(\text{Re } s_0 - d_1)})| > \varepsilon$  then we set  $d_1$  to  $\beta d_1$ , where  $\beta$  is a “jump factor” constant, and repeat until the left end of the interval has integrand no larger than  $\varepsilon$ . The same is done for  $d_2$  on the right end. We find that  $\beta = 1.3$  is a good compromise between making jumps that don’t produce an overly large interval, yet don’t require too many extra integrand evaluations.

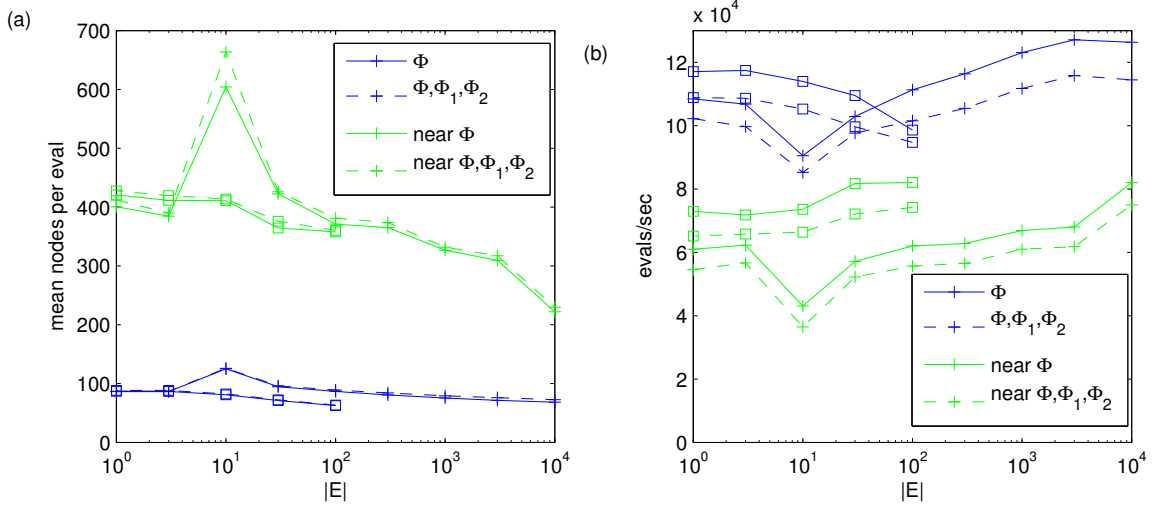


Figure 6: Efficiency of the numerical steepest descent algorithm as a function of frequency parameter  $E$ . (a) Mean number of quadrature nodes used, and (b) mean number of evaluations per second. In both graphs, + signs indicate  $E > 0$  while  $\square$  signs indicate  $E < 0$ . Solid lines are for evaluation of  $\Phi$  alone while dashed lines are for evaluation of  $\Phi$  and its first partials. The source is  $\mathbf{y} = \mathbf{0}$ , and averaging is done over  $10^4$  targets randomly distributed uniformly in angle and uniformly in the logarithm of distance from the origin. For the darker (blue) lines  $|\mathbf{x}| \in [10^{-1}, 10^4]$ , while for the lighter (green) lines only “near” distances  $|\mathbf{x}| \in [10^{-4}, 10^{-1}]$  are used.

The recipe for region A, with saddle points  $s_{\pm}$ , is to use a crude minimization of  $|\psi(e^{\gamma(\alpha)})|$  in  $[\text{Re } s_-, \text{Re } s_+]$ , and if the minimum value exceeds  $\varepsilon$ , to use a single interval  $[\text{Re } s_- - d_1, \text{Re } s_+ + d_2]$ , which is initialized and expanded as before. Otherwise two intervals  $I_1$  and  $I_2$  are used centered at  $s_-$  and  $s_+$  respectively, and each is expanded separately, as before. An example result is shown at the top of Fig. 5.

#### 4.3. Choice of quadrature node spacing

For each interval  $I'$  ( $= I, I_1$  or  $I_2$ ), we need rules to choose  $h$ , the quadrature node spacing in the trapezoid rule approximation to (34),

$$\frac{1}{4\pi} \int_{I'} \exp i\psi(e^{\gamma(\alpha)}) \gamma'(\alpha) d\alpha \approx \frac{h}{4\pi} \sum_{hj \in I'} \exp i\psi(e^{\gamma(hj)}) \gamma'(hj). \quad (38)$$

A general rule is to scale  $h$  in proportion to the minimum *width* of any saddle points contained in  $I'$ . Let  $s_0$  be such a saddle point, then we define its width as

$$\sigma(s_0) := \left| \frac{d^2}{ds^2} \psi(e^s) \right|_{s=s_0}^{-1/2}.$$

Setting a convergence parameter  $h_0$ , we use a node spacing of

$$h = \min \left[ h_{\max}, \frac{|I'|}{n_{\min}}, \sigma h_0 \right]$$

where  $\sigma = \sigma(s_0)$  for the case of one saddle, or  $\sigma = \min[\sigma(s_-), \sigma(s_+)]$  in the case of two. The new numerical parameters here are  $h_{\max}$ , the maximum allowed node spacing, and  $n_{\min}$ , the minimum allowed number of nodes over the interval length  $|I'|$ . Both are needed to prevent  $h$  from become too large, since  $\sigma$  can be arbitrarily large, e.g. when saddles coalesce or when  $|\mathbf{x} - \mathbf{y}|$  is very small.

#### 4.4. Derivatives of $\Phi$

The formula for entries of the matrix approximation to the double-layer operator  $D$  in Sections 3 and (3.1) requires first derivatives of  $\Phi(\mathbf{x}, \mathbf{y})$  with respect to moving the source  $\mathbf{y}$ . These are simple to evaluate from (33)–(34) by passing

the derivative through the integral to give,

$$\frac{\partial \Phi(\mathbf{x}, \mathbf{y})}{\partial y_1} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{-i(x_1 - y_1)e^{-s}}{2} \exp i\psi(e^s) ds \quad (39)$$

$$\frac{\partial \Phi(\mathbf{x}, \mathbf{y})}{\partial y_2} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{-i((x_2 - y_2)e^{-s} - e^s)}{2} \exp i\psi(e^s) ds \quad (40)$$

These may be evaluated with minimal extra effort along with  $\Phi$  by including extra factors in (38). Although these factors can grow exponentially in size, they do not affect the super-exponential decay away from saddle points of the integrand. We take care to include these factors when testing for decay of the integrand to  $\varepsilon$  in Sec. 4.2.

#### 4.5. Convergence and speed tests

We now test the convergence of the above scheme for  $\Phi$  and its derivatives. For true convergence,  $h_0$  must shrink while  $h_{\max}$  also shrinks and  $n_{\min}$  grows; in Fig. 5(b) we perform this test, over the large range of  $E$  and  $\mathbf{x}$  parameters used in Fig. 5(b), 150000 in total. The upper graph shows that a worst-case absolute error around  $3 \times 10^{-11}$  for  $h_0 = 0.35$ ,  $h_{\max} = 0.05$  and  $n_{\min} = 43$ , which we thus find acceptable and fix as our standard choices. In fact, the lower graphs show that typical accuracies are much better, being 13 to 15 digits.

**Remark 5.** *It is known that 24 nodes is sufficient to integrate the Gaussian via the trapezoid rule to double precision accuracy, e.g. [50, Remark 2]. Nearly twice this is needed to guarantee accuracy in our setting, we believe due to distortion around the saddle from an exactly quadratic phase function, and the overshoot in interval size due to  $\beta$  exceeding 1.*

Note that we test *absolute* not *relative* errors in  $\Phi$ : we believe that this is what is relevant for solution of BIEs, and support this claim in the next section. Since  $\Phi$  is exponentially small in the forbidden region, demanding high relative error would require more effort, and is unnecessary.

In Fig. 6(a) we test the mean number of nodes  $n$  used for the contour integral over the test set, splitting the data for near distances  $|\mathbf{x} - \mathbf{y}| < 0.1$ , and for  $|\mathbf{x} - \mathbf{y}| \geq 0.1$ . For the latter, only around 100 nodes is needed, with a slight decrease at large  $E$ . For near distances (hence  $a$  is small), the saddle point  $s_-$  moves leftwards, and the width of the significant region around it grows as shown in Fig. 4(a) and (f). We observe that here  $n$  grows like  $\log 1/|\mathbf{x} - \mathbf{y}|$ . This explains  $n$  in the 200–600 range for near distances. The peak at  $E = 10$  is due to  $I$  being a single large interval containing two saddles, one of which has a small width which demands a small  $h$ .

We implemented the code in C with OpenMP and a MEX interface (constructed via `Mwrap`) to MATLAB (version 2012b), and tested its speed on a desktop workstation with two quad-core Intel Xeon E5-2643 CPUs at 3.3 GHz.<sup>2</sup> Fig. 6(b) shows that at most  $E$  values we achieve a mean rate exceeding  $10^5$  evaluations per second (where we count  $\Phi$  and its two derivatives as a single evaluation). For near distances this drops to around 60% of that. Dips at various  $E$  ranges are explained by the increased  $n$ . The CPU time is believed to be dominated by calls to the complex exponential, and arctangent, functions; memory usage is very small.

## 5. Performance of the boundary value solver

### 5.1. Convergence for interior Dirichlet BVP

To solve the interior BVP corresponding to (1)–(3), firstly the parametrization of the curve  $\partial\Omega$ , and a number  $N$  of boundary nodes, is chosen. Then the data vector  $g_i = -2f(s_i)$ ,  $i = 1, \dots, N$  is filled, and the Nyström matrix  $A$  is filled using (30) for entries away from the diagonal and the Alpert correction of Sec. 3.1 close to the diagonal, with kernel  $K(t, s) = -2(\partial\Phi(\mathbf{z}(t), \mathbf{z}(s))/\partial\mathbf{n}_{\mathbf{z}(s)})|\mathbf{z}'(s)|$ , appropriate for the BIE (23). The dense linear system (29) is solved by direct Gaussian elimination to get the density  $\{\tau_j\}_{j=1}^N$ , and the solution evaluated by direct summation (31).

We test convergence using Dirichlet data  $f = u|_{\partial\Omega}$  coming from the analytic separation of variables solution

$$u(\mathbf{x}) = \cos(\sqrt{E}x_1)\text{Ai}(-x_2), \quad (41)$$

<sup>2</sup>We also tested our codes on a laptop with a quad-core Intel i7-3720QM at 2.6 GHz and found speeds 70%-100% of those reported.



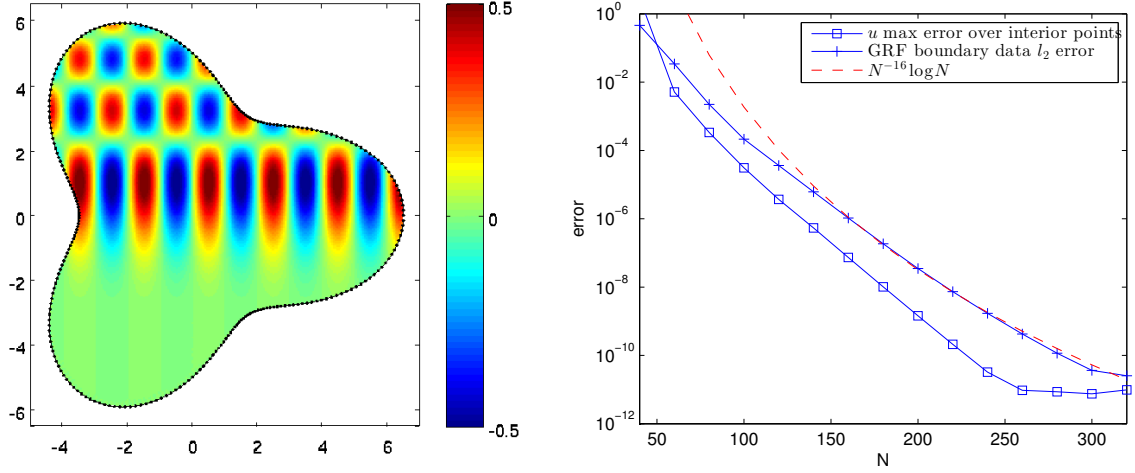


Figure 7: (a) Plot of interior Dirichlet solution  $u$  as evaluated by (31) given the density from solving the BIE, with  $N = 260$  (boundary nodes  $s_i$  shown as dots). (b) Convergence of interior BVP solution: maximum absolute error (□ signs) over 100 interior points chosen randomly to lie inside a copy of  $\partial\Omega$  scaled by 0.8, so points are not too close to  $\partial\Omega$ ; boundary error  $\|S(\partial u/\partial n)^- - (D + \frac{1}{2}I)u^-\|_{l_2}$  (+ signs) for the Green's representation formula with the analytically known data  $\partial u/\partial n$  and  $u$  on  $\partial\Omega$ . See Sec. 5.1.

where  $\text{Ai}$  is the Airy function of the first kind, for  $E = 10$ , with  $\partial\Omega$  a smooth “trefoil” domain given by the polar function  $r(\theta) = 5 + 1.5 \cos(3\theta)$ , about 5 wavelengths across. Fig. 7(a) shows the domain, boundary nodes, and resulting BIE solution constructed via (31). In Fig. 7(b) we observe exponential convergence of the absolute solution error at interior points; we believe this rate is limited by the distance of the nearest points to  $\partial\Omega$  rather than the convergence of the density. At  $N = 260$  we reach 11-digit accuracy (the solution  $u$  has maximum size around 0.5). Filling  $A$  took 5 seconds, and the evaluation of  $u$  at 32841 interior points used to plot Fig. 7(a) took 70 seconds.

As an independent check of the discretization of the operators  $S$  and  $D$  on the boundary, by Green's representation formula [22, (2.5)] [43, Thm. 6.10],

$$u = S(\partial u/\partial n)^- - Du^-, \quad \text{in } \Omega,$$

and thus taking the evaluation point to  $\partial\Omega$  from inside and applying (19)–(20), the boundary function  $S(\partial u/\partial n)^- - (D + \frac{1}{2}I)u^-$  should vanish. We show convergence of its norm in Fig. 7(b); it is consistent with the high order of the Alpert scheme, and reaches 11-digit accuracy (each term, e.g.  $(D + \frac{1}{2}I)u^-$ , has norm 1.5).

$N$	$A$ fill time (s)	dense solve time (s)	evaluation time per target (s)	error
200	3.8	0.004	0.0012	4.1e-05
300	5.3	0.007	0.0018	6.3e-08
400	7.9	0.019	0.0024	2.9e-10
500	10.0	0.023	0.0030	2.6e-12
600	12.6	0.028	0.0036	—

Table 1: Convergence and timing for the small scattering problem shown in Fig. 8 and described in Sec. 5.2. Evaluation time is for the solution  $u$  via (31), and is the mean value over a coarse grid covering the region shown. Error is the maximum absolute error over 100 points lying uniformly on a circle of radius 12 (i.e. a closest distance of 1 from  $\partial\Omega$ ), estimated by comparing to the converged values for  $N = 600$ .

## 5.2. Convergence and timing for scattering problems

For a scattering problem with given incident wave  $u^{\text{inc}}$ , as explained in the introduction, the exterior BVP (1)–(2) is solved with  $f = -u^{\text{inc}}$ . We solve the combined-field BIE (22) similarly to the interior case summarized in Sec. 5.1, except with data  $g_i = 2f(s_i)$ ,  $i = 1, \dots, N$  and kernel  $K(t, s) = 2[\partial\Phi(\mathbf{z}(t), \mathbf{z}(s))/\partial\mathbf{n}_{\mathbf{z}(s)} - i\eta\Phi(\mathbf{z}(t), \mathbf{z}(s))]/|\mathbf{z}'(s)|$ . We test



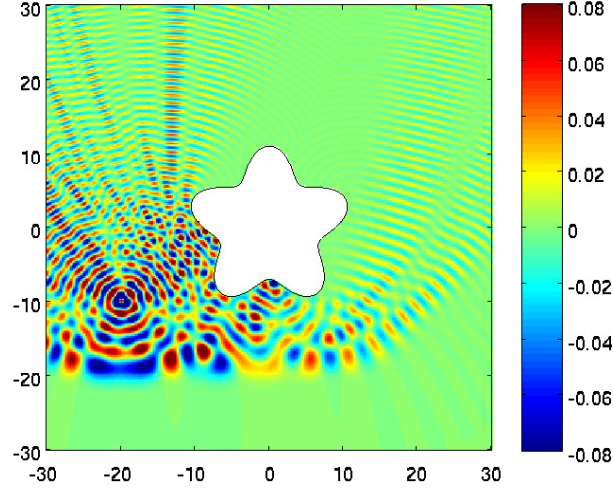


Figure 8: Real part of total wave  $u + u^{\text{inc}}$  for a Dirichlet scattering problem at  $E = 20$ , with  $u^{\text{inc}}(\mathbf{x}) = \Phi(\mathbf{x}, \mathbf{x}_s)$  with  $\mathbf{x}_s = (-20, -10)$ . Around 11 digit accuracy relative to the typical solution size is achieved at  $N = 500$ ; see Sec. 5.2.

$N$	A fill time (s)	dense solve time (s)	evaluation time per target (s)	error
1200	26	0.09	0.008	5.7e-08
1600	37	0.20	0.011	2.1e-10
2000	46	0.31	0.013	2.9e-12
2400	58	0.42	0.016	—

Table 2: Convergence and timing for the large scattering problem shown in Fig. 9 and described in Sec. 5.2. Error is the maximum absolute error over 100 points lying uniformly on a circle of radius 19 (i.e. a closest distance of 1 from  $\partial\Omega$ ), estimated by comparing to the converged values for  $N = 2400$ .

with two smooth scatterers which are chosen to be large enough (diameter of order  $E$ ) that the wavelength has sizeable vertical variation across the object.

We first test a small example, at  $E = 20$ , with shape given by the polar function  $r(\theta) = 9 + 2 \sin(5\theta)$ , which is about 15 wavelengths across at the typical wavenumber  $\sqrt{E}$ . The incident wave is due to a single nearby source at  $\mathbf{x}_s$ . The convergence in Table 1 is consistent with exponential. The solution time is entirely dominated by evaluations of  $\Phi$ , and is consistent with  $10^5$  evaluations per second. The fill time has not yet reached its asymptotic  $O(N^2)$ , since the  $O(30N)$  Alpert correction entries are expensive due to their small source-target distances. The dense linear system solve is  $O(N^3)$ , but insignificant in comparison. A strict  $O(N^2)$  overall scaling is recovered via using an iterative solver; we applied GMRES [51] and found that 43 iterations were required for a residual of  $10^{-12}$ . The total wave solution, shown in Fig. 8, took 4 minutes to evaluate at 84089 grid points, i.e. around 350 target points per second. Notice that the waves bend, and do not propagate below  $x_2 = -E = -20$ .

Finally, we test a similar but more challenging case, at  $E = 65$ , with shape  $r(\theta) = 15 + 3 \cos(10\theta)$ , about 50 wavelengths across. The convergence and timing is in Table 2 and the total wave solution is shown in Fig. 9.<sup>3</sup> Again, 11 digits of accuracy is achieved at  $N = 2000$  (relative to the typical size of  $u$ , which is of order 0.1). For GMRES, 59 iterations were needed to reach a residual of  $10^{-12}$ , showing scarcely any growth from the lower-frequency example. The plot in Fig. 9 took around 50 minutes for 226000 target points, i.e. about 80 target points per second. The parabolic turning point for the source is clearly visible, as well as waves of lower amplitude that have been scattered and hence escape this parabola.

<sup>3</sup>Curiously, fill times on the laptop were slightly faster than for the desktop, but evaluation times were only 70% as fast.

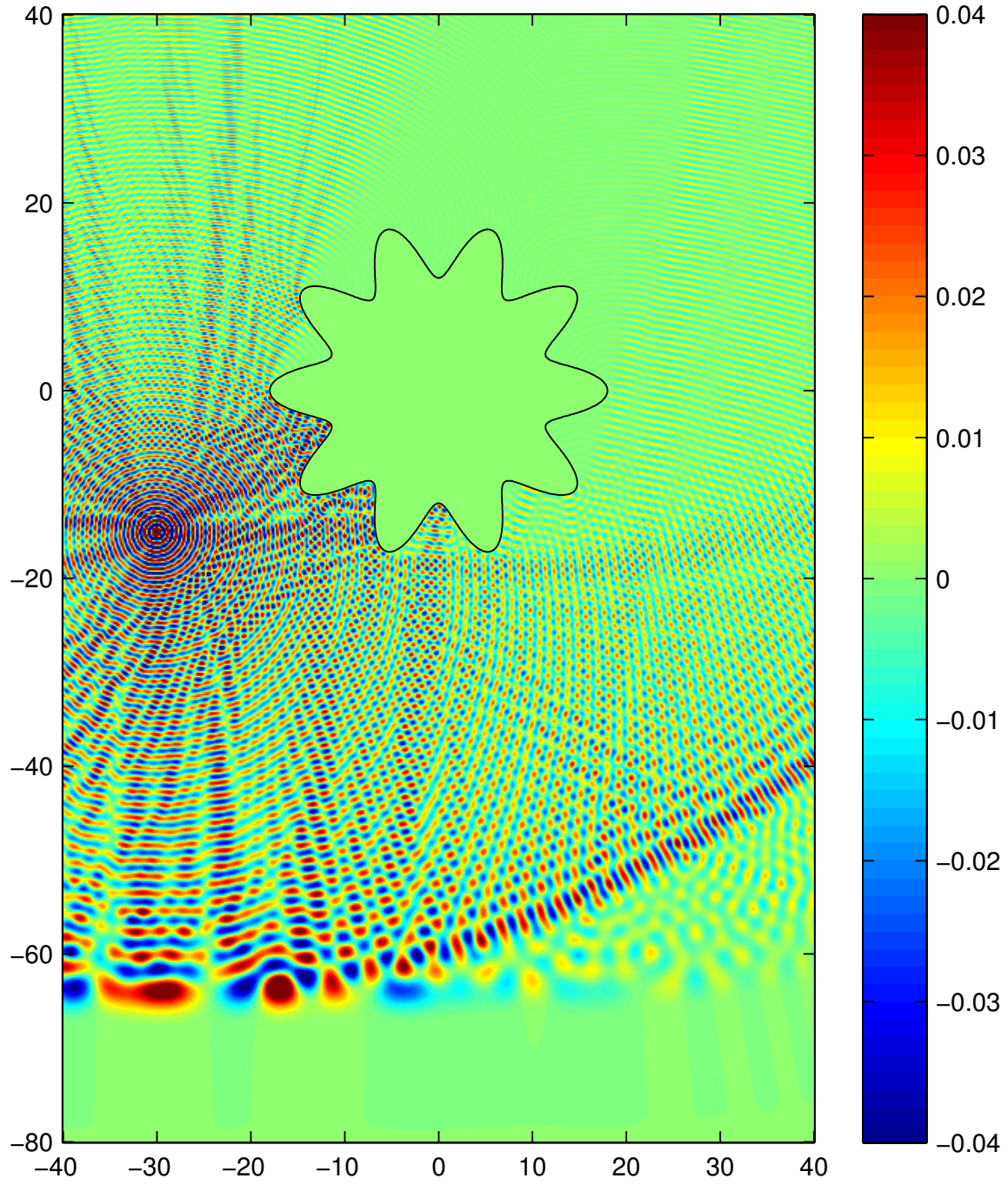


Figure 9: Real part of total wave  $u + u^{\text{inc}}$  for a Dirichlet scattering problem at  $E = 65$ , with  $u^{\text{inc}}(\mathbf{x}) = \Phi(\mathbf{x}, \mathbf{x}_s)$  with  $\mathbf{x}_s = (-30, -15)$ . Around 11 digit accuracy relative to the typical solution size is achieved at  $N = 2000$ ; see Sec. 5.2.

## 6. Conclusion and discussion

We have presented an efficient scheme for high-frequency scattering from smooth objects embedded in a stratified medium in which the inverse square of wave speed varies linearly in the vertical coordinate (the “gravity Helmholtz equation”). Our high efficiency and accuracy comes from combining numerical saddle point integration for an integral representation of the fundamental solution  $\Phi$ , with a boundary integral formulation and high-order quadrature rules for the singular kernels, allowing a problem 50 wavelengths in diameter to be solved to 11 digit accuracy in less than a minute on a desktop or laptop. Our detailed study of the saddle points (and their connection to classical ray dynamics) allows around  $10^5$  evaluations of  $\Phi$  per second, *independent* of the wavenumber. Solution cost is dominated by evaluations of  $\Phi$ , which is trivially parallelizable, and, once the matrix is filled, multiple incident waves at the same  $E$  can be solved with negligible extra cost. The scheme is strictly  $O(N^2)$  when an iterative solver (such as GMRES) is used; here convergence is rapid due to the second-kind formulation.

In addition we placed the boundary value problem in the unbounded stratified medium on a more rigorous footing by deriving radiation conditions (Definition 1) such that the solution is unique. It remains to prove the conjecture that these are indeed satisfied by our causal  $\Phi$ ; this would give an existence proof for the BVP (Remark 3).

In terms of future research, the sound-hard and transmission problems [22] are straightforward variants, as is the restriction to a half-space (reflected rays would need to be considered). The BIE operators we have constructed are also ideal for applying our medium’s radiation boundary conditions to finite-element solvers. When the obstacle is no more than around 100 wavelengths across, much acceleration is possible: a kernel-independent fast multipole method (FMM) [52] could be used to apply  $A$  in each GMRES iteration, or a fast direct solver [26]; both would evaluate only  $O(N)$  as opposed to  $O(N^2)$  matrix elements. The former would also be much faster than direct summation for evaluation of  $u$ . We hope that our numerical saddle point integration techniques might prove useful for other (special) functions. The generalization to 3D will be easy, since  $\Phi$  may be expressed directly using Airy functions [8]. A generalization to *quadratic* variation of the inverse square wave speed is also possible since the time-dependent Schrödinger Green’s function is still known analytically [13]; this could be used for modeling guiding channels in underwater acoustics.

Documented C/OpenMP and MATLAB/MEX codes, with which all tests were performed, are freely available at <http://math.dartmouth.edu/~ahb/software/lhelmfs.tgz>

## Acknowledgements

We have benefited from helpful discussion with Simon Chandler-Wilde, Erik van Erp, and Nick Trefethen. AHB is grateful for support from NSF grant DMS-1216656. BJN is grateful for support from the Paul K. Richter and Evalyn E. Cook Richter Memorial Fund. The work of JMM and BJN was performed while at the Department of Mathematics at Dartmouth College.

## Appendix A. Proof of Theorem 1: uniqueness of radiative solutions

We adapt the radial methods of proof of [22, Thm 3.7] to handle the very different asymptotic behaviors in horizontal and vertical directions. First we need the following Cartesian version of Rellich’s far field decay condition [22, Lemma 2.11].

**Lemma 2** (Cartesian Rellich). *Let  $u$  satisfy (1), with medium (3), in the complement of a bounded domain  $\Omega$ , and*

$$\lim_{x_2 \rightarrow +\infty} k(x_2) \int_{-\infty}^{\infty} |u(x_1, x_2)|^2 dx_1 = 0. \quad (\text{A.1})$$

*Then  $u = 0$  in  $\mathbb{R}^2 \setminus \overline{\Omega}$ .*

*Proof.* For sufficiently large  $x_2$ , using the horizontal Fourier transform  $\hat{u}(\xi, x_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(x_1, x_2) e^{i\xi x_1} dx_1$ , the PDE becomes, for each  $\xi \in \mathbb{R}$ , an ODE in  $x_2$ ,

$$\partial_{x_2}^2 \hat{u}(\xi, x_2) + (x_2 + E - \xi^2) \hat{u}(\xi, x_2) = 0.$$

This is a shifted Airy's equation, thus, in terms of Airy functions Ai and Bi,

$$\hat{u}(\xi, x_2) = \alpha(\xi) \text{Ai}(-x_2 - E + \xi^2) + \beta(\xi) \text{Bi}(-x_2 - E + \xi^2)$$

By unitarity of the Fourier transform, (A.1) implies

$$\lim_{x_2 \rightarrow +\infty} k(x_2) \int_{-\infty}^{\infty} |\hat{u}(\xi, x_2)|^2 d\xi = 0 \quad (\text{A.2})$$

By the asymptotics  $\text{Ai}(-z) \sim \pi^{-1/2} z^{-1/4} \cos(2z^{3/2}/3 - \pi/4)$  and  $\text{Bi}(-z) \sim -\pi^{-1/2} z^{-1/4} \sin(2z^{3/2}/3 - \pi/4)$ , as  $z \rightarrow +\infty$  [53, (9.7.9), (9.7.11)], and  $k(x_2) \sim \sqrt{x_2}$ , it follows that if  $\alpha(\xi)$  or  $\beta(\xi)$  were nonzero on any open subset of  $\mathbb{R}$ , then the limit (A.2) would be positive. Thus  $\alpha$  and  $\beta$  are zero except possibly at a set of measure zero. Taking the inverse Fourier transform,  $u(x_1, x_2) = 0$  for all  $x_2$  sufficiently large. Since (1) has analytic coefficients, its solutions are analytic in both variables. By unique continuation,  $u = 0$  in all of  $\mathbb{R}^2 \setminus \bar{\Omega}$ .  $\square$

Next we need flux conservation, which states that, for any bounded region  $D \subset \mathbb{R}^2$  with boundary  $\partial D$  in which  $u$  satisfies (1) with  $k(x_2)^2$  real,

$$-\text{Im} \int_{\partial D} u \bar{u}_n ds = 0, \quad (\text{A.3})$$

where  $u_n = \partial u / \partial n$  is the outward-pointing normal derivative. The left-hand side may be interpreted as the wave energy flux exiting the domain  $D$ . This follows simply from taking the imaginary part of Green's first identity

$$\int_{\partial D} u \bar{u}_n ds = \int_D u \Delta \bar{u} + |\nabla u|^2 d\mathbf{x}$$

after inserting  $\Delta u = -k(x_2)^2 u$  from the PDE.

We now prove a result analogous to [22, Thm. 2.12].

**Theorem 2** (Non-negative incoming flux). *Let  $\Omega \subset \mathbb{R}^2$  be a bounded domain. Let  $u$  solve (1) with medium (3) in  $\mathbb{R}^2 \setminus \bar{\Omega}$ , be radiative according to Definition 1, and have non-negative incoming flux, i.e.,*

$$\text{Im} \int_{\partial \Omega} u \bar{u}_n ds \geq 0.$$

*Then  $u = 0$  in  $\mathbb{R}^2 \setminus \bar{\Omega}$ .*

*Proof.* Expanding the square in (5) gives

$$\lim_{x_2 \rightarrow +\infty} \lim_{M \rightarrow \infty} \int_{-M}^M \frac{1}{k(x_2)} \left| \frac{\partial u}{\partial x_2} \right|^2 + k(x_2) |u|^2 dx_1 + 2 \text{Im} \int_{-M}^M u \frac{\partial \bar{u}}{\partial x_2} dx_1 = 0.$$

Applying (A.3) to the punctured rectangle  $(-M, M) \times (-x_2, x_2) \setminus \bar{\Omega}$ , by the decay conditions (6)–(7) and Cauchy-Schwarz the flux contributions from the bottom, left, and right sides vanish, giving

$$\lim_{x_2 \rightarrow +\infty} \lim_{M \rightarrow \infty} \int_{-M}^M \frac{1}{k(x_2)} \left| \frac{\partial u}{\partial x_2} \right|^2 + k(x_2) |u|^2 dx_1 = -2 \text{Im} \int_{\partial D} u \bar{u}_n ds$$

analogous to [22, (2.10)]. By the assumption of the theorem, the right-hand side is non-positive, so (A.1) holds, and Lemma 2 completes the proof.  $\square$

Finally, to prove the uniqueness of the radiative solution to the Dirichlet BVP (1)–(2), we need only that if  $u = 0$  on  $\partial \Omega$ , and  $u$  is a radiative solution, then  $u = 0$  in  $\mathbb{R}^2 \setminus \bar{\Omega}$ . Given the remark in the proof [22, Thm 3.7] about the convergence of the normal derivative, the incoming flux is zero and the result follows from Theorem 2. We suspect that the above generalizes easily to more general profiles  $k(x_2)$ .

## References

- [1] J. B. Keller, J. S. Papadakis (Eds.), *Proceedings of the 1974 Workshop on Wave Propagation and Underwater Acoustics*, Lecture Notes in Physics, 70, Springer-Verlag, 1977.
- [2] P. C. Etter, *Underwater acoustic modeling and simulation*, 4th Edition, CRC Press, 2013.
- [3] E. Premat, Y. Gabillet, A new boundary-element method for predicting outdoor sound propagation and application to the case of a sound barrier in the presence of downwards refraction, *J. Acoust. Soc. Am.* 108 (6) (2000) 2775–2783.
- [4] D. R. Hartree, Optical and equivalent paths in a stratified medium, treated from a wave standpoint, *Proc. Roy. Soc. Lond. A* 131 (817) (1931) 428–450.
- [5] A. B. Shvartsburg, V. Kuzmiak, G. Petite, Optics of subwavelength gradient nanofilms, *Phys. Rep.* 452 (2007) 33–88.
- [6] C. H. Chapman, *Fundamentals of Seismic Wave Propagation*, Cambridge University Press, 2004.
- [7] S. Suresh, A. Mortensen, *Fundamentals of functionally graded materials*, Maney Materials Science, 1998.
- [8] C. Bracher, W. Becker, S. A. Gurvitz, M. Kleber, M. S. Marinov, Three-dimensional tunneling in quantum ballistic motion, *Am. J. Phys.* 66 (1998) 38–48.
- [9] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, Wiley-IEEE Press, 1999.
- [10] B. Gottlieb, M. Kleber, J. Krause, Tunneling from a 3-dimensional quantum well in an electric field: an analytical solution, *Z. Phys. A* 339 (1991) 201–206.
- [11] R. Kress, On the numerical solution of a hypersingular integral equation in scattering theory, *J. Comput. Appl. Math.* 61 (1995) 345–360.
- [12] D. Marcuse, Influence of curvature on the losses of doubly clad fibers, *Appl. Optics* 21 (23) (1982) 4208–4213.
- [13] E. J. Heller, Wavepacket dynamics and quantum chaos, in: *Chaos et physique quantique* (Les Houches, 1989), North-Holland, Amsterdam, 1991, pp. 547–664.
- [14] A. Gil, J. Segura, N. M. Temme, *Numerical Methods for Special Functions*, SIAM, 2007.
- [15] L. N. Trefethen, J. A. C. Weideman, The exponentially convergent trapezoidal rule, *SIAM Review* 56 (3) (2014) 385–458.
- [16] R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations*, SIAM, 2007.
- [17] I. M. Babuska, S. A. Sauter, Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?, *SIAM J. Numer. Anal.* 34 (6) (1997) 2392–2423.
- [18] B. Engquist, L. Ying, Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers, *Multiscale Mod. Sim.* 9 (2) (2011) 686–710.
- [19] B. Engquist, A. Majda, Absorbing boundary conditions for the numerical simulation of waves, *Math. Comp.* 31 (1977) 629–651.
- [20] F. Collino, P. Joly, Splitting of operators, alternate directions, and paraxial approximations for the three-dimensional wave equation, *SIAM J. Sci. Comput.* 16 (5) (1995) 1019–1048.
- [21] N. A. Kampanis, V. A. Dougalis, J. A. Ekaterinaris (Eds.), *Effective computational methods for wave propagation*, CRC Press, Boca Raton, 2007.
- [22] D. Colton, R. Kress, *Inverse acoustic and electromagnetic scattering theory*, 2nd Edition, Vol. 93 of Applied Mathematical Sciences, Springer-Verlag, Berlin, 1998.
- [23] K. Atkinson, *The numerical solution of integral equations of the second kind*, Cambridge University Press, 1997.
- [24] W. Y. Crutchfield, Z. Gimbutas, G. L., J. Huang, V. Rokhlin, N. Yarvin, J. Zhao, Remarks on the implementation of the wideband FMM for the Helmholtz equation in two dimensions, Vol. 408 of *Contemp. Math.*, Amer. Math. Soc., Providence, RI, 2006, pp. 99–110.
- [25] W. Hackbusch, A sparse matrix arithmetic based on H-matrices; Part I: Introduction to H-matrices, *Computing* 62 (1999) 89–108.
- [26] A. Gillman, P. Young, P. Martinsson, A direct solver with  $O(N)$  complexity for integral equations on one-dimensional domains, *Frontiers of Mathematics in China* 7 (2) (2012) 217–247.
- [27] R. Kress, Boundary integral equations in time-harmonic acoustic scattering, *Mathl. Comput. Modelling* 15 (1991) 229–243.
- [28] S. Hao, A. H. Barnett, P. G. Martinsson, P. Young, High-order accurate Nyström discretization of integral equations with weakly singular kernels on smooth curves in the plane, *Adv. Comput. Math.* 40 (1) (2014) 245–272.
- [29] P. G. Martinsson, V. Rokhlin, A fast direct solver for scattering problems involving elongated structures, *J. Comput. Phys.* 221 (2007) 288–302.
- [30] A. Kirsch, P. Monk, An analysis of the coupling of finite-element and Nyström methods in acoustic scattering, *IMA J. Numer. Anal.* 14 (1994) 523–544.
- [31] A. Gillman, A. H. Barnett, P.-G. Martinsson, A spectrally accurate direct solution technique for frequency-domain scattering problems with variable media, <http://arxiv.org/abs/1308.5998>, to appear, *BIT J. Numer. Math.*, 2014.
- [32] M. H. Cho, W. Cai, A parallel fast algorithm for computing the Helmholtz integral operator in 3-D layered media, *J. Comput. Phys.* 231 (2012) 5910–25.
- [33] C. H. Wilcox, *Sound propagation in stratified fluids*, Applied Mathematical Sciences, volume 50, Springer-Verlag, 1984.
- [34] C. Jeres-Hanckes, J.-C. Nédélec, Asymptotics for Helmholtz and Maxwell solutions in 3-D open waveguides, *Commun. Comput. Phys.* 11 (2) (2012) 629–646.
- [35] W. L. Miranker, The reduced wave equation in a medium with variable index of refraction, *Comm. Pure Appl. Math.* 10 (1957) 491–502.
- [36] S. N. Chandler-Wilde, Boundary value problems for the Helmholtz equation in a half-plane, in: *Proceedings of the 3rd International Conference on Mathematical and Numerical Aspects of Wave Propagation*, Mandelieu-La Napoule, France, April 1995, SIAM, 1995, pp. 188–197.
- [37] G. N. Watson, *A Treatise on the Theory of Bessel functions*, Cambridge University Press, 1922.
- [38] P. R. Garabedian, *Partial differential equations*, John Wiley & Sons Inc., New York, 1964.
- [39] J. B. Keller, Progress and prospects in the theory of linear wave propagation, *SIAM Review* 21 (2) (1979) 229–245.
- [40] L. C. Evans, *Partial Differential Equations*, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 1998.
- [41] G. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective*, University Science Books, 2007.
- [42] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics, AMS, 2012.

- [43] W. C. H. McLean, Strongly elliptic systems and boundary integral equations, Cambridge University Press, 2000.
- [44] R. Kress, Numerical Analysis, Graduate Texts in Mathematics #181, Springer-Verlag, 1998.
- [45] B. K. Alpert, Hybrid Gauss-trapezoidal quadrature rules, SIAM J. Sci. Comput. 20 (1999) 1551–1584.
- [46] R. Kress, Linear Integral Equations, 2nd Edition, Vol. 82 of Appl. Math. Sci., Springer, 1999.
- [47] A. H. Barnett, Evaluation of layer potentials close to the boundary for Laplace and Helmholtz problems on analytic planar domains, SIAM J. Sci. Comput. 36 (2) (2014) A427–A451.
- [48] F. Bornemann, G. Wechsberg, Optimal contours for high-order derivatives, IMA J. Numer. Anal. 33 (2013) 403–412.
- [49] D. Huybrechs, S. Vandewalle, On the evaluation of highly oscillatory integrals by analytic continuation, SIAM J. Numer. Anal. 44 (2006) 1026–1048.
- [50] L. Greengard, J.-Y. Lee, Accelerating the nonuniform fast fourier transform, SIAM Review 46 (3) (2004) 443–454.
- [51] Y. Saad, M. H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Stat. Sci. Comput. 7 (3) (1986) 856–869.
- [52] L. Ying, G. Biros, D. Zorin, A kernel-independent adaptive fast multipole method in two and three dimensions, J. Comput. Phys. 196 (2) (2004) 591–626.
- [53] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, C. W. Clark (Eds.), NIST Handbook of Mathematical Functions, Cambridge University Press, 2010, <http://dlmf.nist.gov>.